

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2001, 16:07:43 ; Search time 16.07 seconds
(without alignments)
1602.178 Million cell updates/sec

Title: US-09-633-145-2

Sequence: 1 MPFCHNINISCVKNNMSND.....FGRIPOKDSRCKLFLELSS 338

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR_68:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 677.5 | 37.6 | 337 | 2 | JC5832 |
| 2 | 529 | 29.4 | 387 | 2 | S55550 |
| 3 | 529 | 29.4 | 406 | 2 | S55849 |
| 4 | 510.5 | 28.3 | 459 | 2 | A56849 |
| 5 | 497.5 | 27.6 | 465 | 2 | I51661 |
| 6 | 493.5 | 27.4 | 363 | 2 | I50475 |
| 7 | 486 | 27.0 | 358 | 2 | J01278 |
| 8 | 482 | 26.8 | 444 | 2 | C55886 |
| 9 | 482 | 26.8 | 450 | 2 | A55886 |
| 10 | 481 | 26.7 | 451 | 2 | I51660 |
| 11 | 480 | 26.7 | 451 | 2 | I51659 |
| 12 | 475 | 26.4 | 377 | 2 | S68423 |
| 13 | 474 | 26.3 | 377 | 2 | A53279 |
| 14 | 473 | 26.3 | 374 | 2 | I77467 |
| 15 | 472.5 | 26.2 | 359 | 2 | JH0449 |
| 16 | 467 | 25.9 | 446 | 2 | I47217 |
| 17 | 466.5 | 25.9 | 377 | 2 | B30341 |
| 18 | 466.5 | 25.9 | 366 | 2 | S72168 |
| 19 | 462.5 | 25.7 | 463 | 2 | B56849 |
| 20 | 461 | 25.6 | 433 | 2 | A25896 |
| 21 | 459.5 | 25.5 | 359 | 2 | JC4120 |
| 22 | 456 | 25.4 | 446 | 1 | DYH0D1 |
| 23 | 454.5 | 25.2 | 477 | 1 | DYH0D5 |
| 24 | 452 | 25.1 | 418 | 1 | QRAYB2 |
| 25 | 451 | 25.0 | 418 | 2 | S10855 |
| 26 | 450.5 | 25.0 | 359 | 2 | A43908 |
| 27 | 450 | 25.0 | 475 | 2 | A41271 |
| 28 | 449.5 | 25.0 | 464 | 2 | S12591 |
| 29 | 449 | 24.9 | 418 | 2 | S00260 |

| | | | | | | |
|----|-------|------|-----|---|--------|---------------------|
| 30 | 446.5 | 24.8 | 415 | 2 | I53040 | beta-2 adrenergic |
| 31 | 446.5 | 24.8 | 487 | 1 | DYRTD1 | dopamine receptor |
| 32 | 446 | 24.8 | 466 | 2 | S56794 | beta-1-adrenergic |
| 33 | 438.5 | 24.3 | 413 | 1 | QRH0B2 | beta-2-adrenergic |
| 34 | 437.5 | 24.3 | 486 | 2 | B55886 | dopamine receptor |
| 35 | 433.5 | 24.1 | 477 | 1 | QRH0B1 | beta-1-adrenergic |
| 36 | 433.5 | 24.0 | 366 | 2 | A47321 | serotonin receptor |
| 37 | 429.5 | 23.8 | 428 | 2 | A55044 | beta-4C-adrenergic |
| 38 | 429 | 23.8 | 480 | 2 | I53053 | beta-1-adrenergic |
| 39 | 427.5 | 23.7 | 366 | 2 | S26048 | serotonin receptor |
| 40 | 427.5 | 23.7 | 386 | 2 | A42688 | serotonin receptor |
| 41 | 427.5 | 23.7 | 386 | 2 | S18637 | serotonin receptor |
| 42 | 427 | 23.7 | 477 | 2 | S71323 | serotonin receptor |
| 43 | 425.5 | 23.6 | 429 | 2 | S65656 | alpha-1A adrenergic |
| 44 | 425.5 | 23.6 | 466 | 2 | JN0765 | alpha-1C-adrenergic |
| 45 | 425.5 | 23.6 | 499 | 2 | S05657 | alpha-1C-adrenergic |

ALIGNMENTS

RESULT 1

JC5832
neurotransmitter receptor - human
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1998 #sequence_revision 13-Mar-1998 #taxid_change 24-Nov-1999
C:Accession: JC5832
R:Zeng, Z.; Fan, P.; Rand, E.; Kyaw, H.; Su, K.; Madhve, V.; Carter, K.C.; Li, Y.
Biochem. Biophys. Res. Commun. 242, 575-578, 1998
A:Title: Cloning of a putative human neurotransmitter receptor expressed in skeletal
A:Reference number: JC5832; MUID:98125534
A:Accession: JC5832
A:Molecule type: mRNA
A:Residues: 1-337 <ZEN>
A:Cross-references: GB:AF021818; NID:92465431; PIDN:AAC30581.1; PID:92465432
A:Experimental source: brain
A:Genetics:
A:Map position: 6q23
C:Superfamily: vertebrate rhodopsin
C:Keywords: glycoprotein; receptor; transmembrane protein
F:29-63/Domain: transmembrane #status predicted <TM1>
F:66-95/Domain: transmembrane #status predicted <TM2>
F:114-118/Domain: transmembrane #status predicted <TM3>
F:149-173/Domain: transmembrane #status predicted <TM4>
F:186-200/Domain: transmembrane #status predicted <TM5>
F:204-229/Domain: transmembrane #status predicted <TM6>
F:253-310/Domain: transmembrane #status predicted <TM7>
F:21/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 37.6%; Score 677.5; DB 2; length 337;
Best local Similarity 39.0%; Pred. No. 3.2e-49;

Matches 130; Conservative 67; Mismatches 125; Indels 11; Gaps 4;

| | | | | | |
|----|-----|---|----|-----|--|
| QY | 3 | FCHNINISCVKNNMSNDVRSALSLMVLITLTNGNLIVISISFQJLHTPTMMLH 62 | DB | 16 | FCYO-VNGSCPRVHTLGIQLVITLCAACMLITVLGNVFAVAVSFAKLRPTNELL 74 |
| QY | 63 | SMATVDFLLCLVMPYSMVRSAEHCWFEVGEVCKIHSTDMSSASIFPLSTISIDRY 122 | DB | 75 | SLALADFLGLLPLSTINSVSCNFFGDFLCRLHTLDTLCLASIFPLCFISIDRRC 134 |
| QY | 123 | AVCDPLRYKAKNNILVYCVMIFTSWSPVAFAGMFLNLFAGAEIYKHHGCGCS 182 | DB | 135 | AICDDPLYPSPKFTVVALRYLLAGMGVPAVYTSLEFYTDVETRLSQ-MLEMRPCVGSQ 193 |
| QY | 183 | VFSKISGVITFTFSYRISIMLCVYRIRYLAKQARLISDAWQLOTGLBKGISQ 242 | DB | 194 | LILNKRWGWLNF-PLFVFPCLMISLYVKIFVATROAOQITFLCGL-----AGAA 244 |
| QY | 243 | SKERKAVKTLGIWGVLLCWCFFICTVADPLHIIPTLNVYVFGYLNSTENPAV 302 | DB | 245 | KHERKAKTKLGIWGVLLCWLMPFTITDVTWDSLHRTTPPLVFDLFIWAFNSACNPIT 304 |

```

Oy      303 YAFYPMFRKALKMLFGKIFQKDSRCLE 335
        | | | | | | | | | | | | | | | |
Db      305 YFSYQMFRRKALKLILSKVPSPTRTVDLQGE 337

RESULT      2
S55550
5-Ht4S receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 26-Aug-1999
C:Accession: S55550
R:Genalt, C.; Adham, N.; Kao, H.T.; Olsen, M.A.; Laz, T.M.; Schechter, L.E.; Bard, J.A.;
EMBO J. 14, 2806-2813, 1995
A:Title: The 5-HT(4) receptor: molecular cloning and pharmacological characterization of
A:Reference number: S55549; MUID:95317299
A:Accession: S55550
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-387 <GPR>
A:Cross-references: GB:U20906; NID:g924638; PIDN:AMC52232.1; PID:g924639
C:Superfamily: vertebrate rhodopsin

```

```

Query Match      29.4%; Score 529; DB 2; Length 387;
Best Local Similarity    35.4%; Pred. No. 8.9e-37;
Matches 111; Conservative   61; Mismatches 114; Indels   28; Gaps

```

| | | | |
|----|-----|--|-----|
| OY | 25 | IYSLWLIITLTVLNLYIVSISHPKQLH-TPTNMLIHSMAVDFLLGCLVPMSMVS | 83 |
| | | : : : : : : : : : : : : : : | |
| Dd | 22 | LITFPAMYLIMLENLVMVAVCBDRQLRKTXTNFIVLSLAFADLVLVNAFCAL | 81 |
| OY | 84 | AECWFVEGEVCKIHSTSDIMLSASIFHLSTSIDRYAV-CDPLRYAKKNILVICYM | 142 |
| | | : : : : : : : : : : : : : : | |
| Dd | 83 | VODIFYEGEMCFVLRTSLDVLTLTTSIFHLCSIDRYAIIQQPLVYNKKMTPLEIAM | 141 |
| OY | 143 | ITISSVPAVFAGMFELFNKGAEIY----YKHVCRGGCVSPFSGVISGVLFMTFS | 198 |
| | | : : : : : : : : : : : : : : | |
| Dd | 142 | DGGCWIPPIFSFLPMOGMNINIGYDIIEKRKFHNNSSTFCVFWVNKPATITCSVAF | 201 |
| OY | 199 | YTPGISIMLCVVYRIYLIAEQARLISDAOQLOLGLEMNGISQSK----- | 244 |
| | | : : : : : : : : : : : : : : | |
| Dd | 202 | YIPELLMVATYYRIYTAKEHA-----QQIQDM-LQRGAGVSESPPADQSHTRMR | 253 |
| OY | 245 | EKKAVYTCIGVWGVELICMCFFICTVMDPFLHYIIPPLINDVLIIFGLNSTFENMYVA | 304 |
| | | : : : : : : : : : : : : : : | |
| Dd | 254 | EKKAAETLCVINGCFCFCWAPFEVTNIVDPFIDYTVPEKYMTAFELMGYINSGLNPFLYA | 313 |
| OY | 305 | FEPYPRFRALKMML 318 | |
| | | : : : : : : : : : : : : : : | |
| Dd | 314 | FLNKSFRRAPFLIIL 327 | |

RESULT 3
 S55549
 serotonin 4 receptor - rat
 N:Alternate names: 5-HT receptor 4L
 C:Species: *Rattus norvegicus* (Norway rat)
 C:date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 24-Nov-1995
 C:Accession: S55549; S66494
 R:Rigler, C.; Agham, N.; Kao, H.T.; Olsen, M.A.; Laz, T.M.; Schechter, L.E.; Bard, J.A.;
 EMO J. 14, 2806-2815, 1995
 A:title: The 5-HT(4) receptor: molecular cloning and pharmacological characterization of
 A:Reference number: S55549; MUID:95317299
 A:Accession: S55549
 A:status: preliminary
 A:molecule type: mRNA
 A:residues: 1-406 <GCR>
 A:cross-references: GB:U020907; NID:9924640; PIDD:AA052233.1; PID:9924641
 R:Ullmer, C.; Schmuck, K.; Kalkman, H.O.; Luebbert, H.
 FEBS Lett. 370, 215-221, 1995
 A:title: Expression of serotonin receptor mRNAs in blood vessels.
 A:Reference number: S66487; MUID:95385798

A:Accession:566494
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 95-259
A:Cross-references: EMBL:Z48153; MID:9984171; PIDD:CAA8170.1; PID:9984172
A:Experimental source: tissue brain
A:Superfamily: vertebrate rhodopsin
C:Keywords: neurotransmitter receptor

[illegible]

RESULT 4
A56849 dopamine receptor-like protein D14 - Japanese pufferfish
C:Species: Fugu rubripes (Japanese pufferfish)
C:Date: 27-Oct-1995 #sequence_revision 27-Oct-1995 #text_change 20-Jun-2000
C:Accession: A56849
R:Macrae, A.D.; Brenner, S.
Genomics 25, 436-446, 1995
A:Title: Analysis of the dopamine receptor family in the compact genome of the puffer
A:Reference number: A56849; WUID:95309911
A:Accession: A56849
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-459 <MAC>
A:Cross-references: GB:X80174; NID:g1204089; PIDN:CAA56455.1; PID:g1204090
C:Superfamily: vertebrate rhodopsin
C:Keywords: neurotransmitter receptor

[illegible]

| | | | |
|---------------------------------|---|---|-----|
| Oy | | PGSMILCYVRYRIYLIAKQARLI-----SDANOKLOIGLEMKNGISOSKER | 246 |
| | | I::I:: I::I::IIAKOIRISALENAAESAOQRHSSMGNSLSMESDCSFKNMFRET | 266 |
| Dd | 207 | PVALIVITYYTHIYRKAQOIKRISALENAAESAOQRHSSMGNSLSMESDCSFKNMFRET | 266 |
| Oy | 247 | KAVKTLCIVMCGVFLLCMCPFFCTGVDPDLHY-----IIPPLNDVLIMFGYLNSTFN | 299 |
| | | I::I::I::I::III:: III:: II:: I::I::I::I::I::I:: | |
| Dd | 267 | KVKLTLSVMGCVOCMLPEFILNCMVPFCEADDTTDEPCISSTFEDVFWNKGMAANSLSL N | 326 |
| Oy | 300 | PMVTAEFYFWRKKALKMM L 318 I::I::I:: III:: ::I | |
| Dd | 327 | PIYAFAFD-N-FRKAFSILL 344 | |
| RESULT | 5 | | |
| | 151661 | dopamine D1C receptor - African clawed frog (fragment) | |
| C.Species: | Xenopus laevis (African clawed frog) | | |
| C.Date: | 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 13-Aug-1999 | | |
| C.Accession: | F51661 | | |
| R.Sugamori, | K.S.; Demchishyn, L.L.; Chung, M.H.; Miznik, H.B. | | |
| A.Proc. Natl. Acad. Sci. U.S.A. | 91, 10536-10540, 1994 | | |
| A.Title: | D1A, D1B, and D1C dopamine receptors from Xenopus laevis. | | |
| A.Reference number: | F51659; MUID:95024150 | | |
| A.Accession: | F51661 | | |
| A.Status: | preliminary; translated from GB/EMBL/DDBU | | |
| A.Molecule type: | DNA | | |
| A.Residues: | 1-465 <SD> | | |
| C.Cross-references: | EMBL,U07865; NID:g559763; PID:NAAA50830.1; PUD:g559764 | | |
| C.Superfamily: | vertebrate rhodopsin | | |
| C.Keywords: | neurotransmitter receptor | | |

```

Query Match          27.6%; Score 497.5; DB 2; Length 465;
Best Local Similarity 34.5%; Pred. No. 4.5e-34;
Matches 118; Conservative 61; Mismatches 130; Indels 33; Gaps 8;

QY      6 NIINISCVNNMSND-----VRASLYSLMVLIIITTEVGNLIVYSISHFKQLATP 56
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      5 STEFNVT--NWVHADDDVGNSDLSLRALTGLLSLIILSTLTGNTLVCLAVKFRRLRSK 62
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      57 -TNMLISHMATVDPLLCGLMPYSHVRSAEHCWFGVEFCKLHTSTDIMLSSASIFPHSE 115
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      63 VNFNFISSLAVSDFLLVALLVPMKAVEAGFWEGD-FCDTWVAFDIMCSPASILNCI 121
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      116 ISIDRYAVCDPLRKAKANNILVICMFISMSVPAAVFEGFIETLNKGADIELYYKHV 175
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      122 ISIDRWAIASPFERRERKMTORVAFIMIGVAMTLLISLISFIVOLSMHSHEADELNCV 181
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      176 HCRGGGSVFESKISGLTLPTMSFYIPGSIIMLCVYRIYLIAEQARLLISDN-----OK 229
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      182 NHTENDDSSLNRTYAISLSSLSIFYIPVVMIGYIRIYIAOTQIRISSLERAVEHNAQR 241
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      230 LOIGLEMKNGISOS--KERKAVKTGIYGVGFICMCPEFICVMDPPFLHYIIP----- 281
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      242 CSRSLSENESLKTSFPERKRVKLKTLISIMGVFFCMLPFVYNCHLPFGHMMLPGQNEPE 301
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      282 -----PTLNDVLIMEGYLSTFNPVYAFEPYFRKALKMML 318
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      302 PCGVSETTFENIFVWFECMANSSLNPVIYAFNAD-FRKAFTTIL 342
       : |::| | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT    6
150475
dopamine D1 receptor - goldfish
C:Species: Carassius auratus (goldfish)
C:date: 13-Sep-1996 #sequence-revision 13-Sep-1996 #text-change 13-Aug-1999
C:Accession: I50475
R:Fall, D.E.; Manelli, A.M.; Witte, D.G.; Lin, C.W.; Steffey, M.E.; Mackenzie, R.G.
Mol. Pharmacol. 44, 1113-1118, 1993
A:title: Cloning and characterization of a truncated dopamine D1 receptor from goldfish
A:Reference number: I50475; MID:94088471
A:Accession: I50475
```

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-363 <PRA>
A: Cross-references: GB:J08602; NID: g212949; PIDN: AAA16322.1; PID: g212950
A: Superfamily: Vertebrate rhodopsin
C: Keywords: neurotransmitter receptor

[illegible]

```

RESULT 7
histamine H2 receptor - rat
J01278
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Nov-1999
C:Accession: J01278
R:Ruat, M.; Traiffort, E.; Arrang, J.M.; Leurs, R.; Schwartz, J.C.
Biochem. Biophys. Res. Commun. 179, 1470-1478, 1991
A:Title: Cloning and tissue expression of a rat histamine H2-receptor gene.
A:Reference number: J01278; MUID:92028890
A:Accession: J01278
A:Molecule type: DNA
A:Residues: 1-358 <R0>
A:Cross-references: GB:557565; MID:9236183; PIDN:AM19935;1; PID:9236184
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane P
F:2-45/Domain: transmembrane #status predicted <TM1>
F:58-81/Domain: transmembrane #status predicted <TM2>
F:93-113/Domain: transmembrane #status predicted <TM3>
F:136-159/Domain: transmembrane #status predicted <TM4>
F:178-203/Domain: transmembrane #status predicted <TM5>
F:234-260/Domain: transmembrane #status predicted <TM6>
F:267-288/Domain: transmembrane #status predicted <TM7>
F:4/Binding site: carboxylate (Asn) (covalent) #status predicted
F:220,311,315/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 27.0% Score 486; DB 2: Length 358;
Best Local Similarity 34.0% Pred No.3.2e-33;
Matches 117; Conservative 56; Mismatches 123; Indels 48; Gaps 7;

11 SCVKNMSNDVRASLSTLWVLIITLWGNLIYVISHFQKQIHTPTNMLHSMATYDFL 70
11 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 9 SCCLDSMAUKVITSV--VLTFTILITAGVNVVCAVSLNRRLRSITNCFIVSLATLDL 66
11 LGCLVWPYMSVAEHCWTFGEVEFCIKHTSTDMISSAIFHLSPISIDRYVAVDPLRY 130

```

```

Db 67 LGLVLPESAIVQISFTWSFGHVCNITYTSLDVMCTASINLEMISIDRYACVADPLKY 126
Qy 131 KAKMNIIVICVMIFISMSVPAVAFGMIFLELNFGKAEIYYKHVHCRG-----CSVFF 185
Db 127 PVLVTPVRAIVSLVITWISITLSLSTLHGMNSRNG-----TRGNDTPFKQGV 177
Qy 166 SKISGVLFPMISFYIPGSIIMLCVYRIRYILIAKEQARLISDANOKIQIGLEMKNGISQSK- 244
Db 178 NEVYGLVQGLVTFYLPFLIMCVTYRIEIKIAREQAKRI-----NHISWKA 223
Qy 245 -----ERKAKVTGIYMGVFLICMCPFFICTVMDPLHY-----IIPPLNDVLIWFGY 293
Db 224 ATRERKATVITLAAVAGAFICMFPFTAFV-----YGLGDDAIDNAAVEGIYMLGY 277
Qy 294 INSTENPMVYAFEPYFWRKALKMMLGKIFQKDSRCKLFEELS 337
Db 278 AASALNPILYALNDRFRAVYQQLFHCCKFASHNSHRTSLRNN 321

```

```

RESULT 8
C55886
dopamine receptor D1D - chicken
C:Species: Gallus gallus (chicken)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 17-Mar-1999
C:Accession: C55886
R:Demichyshyn, L.L.; Sugamori, K.S.; Lee, F.J.S.; Hamadanizadeh, S.A.; Niznik, H.B.
J. Biol. Chem. 270, 4005-4012, 1995
A:Title: The dopamine D1D receptor. Cloning and characterization of three pharmacologic
A:Reference number: A55886; MUID:95181368
A:Accession: C55886
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-444 <DEM>
A:Cross-references: GB:L336879
C:Superfamily: vertebrate rhodopsin
C:Keywords: neurotransmitter receptor

```

```

Query Match 26.8%; Score 482; DB 2; Length 444;
Best Local Similarity 34.7%; Pred. No. 8.4e-33;
Matches 112; Conservative 57; Mismatches 114; Indels 40; Gaps 8;

Qy 21 VASLSTLWVLIITLVNLIYVISHFKQLHP-TNMLIHSMATVDFLGLVMPYS 79
Db 11 LRLAGCLGALVGLAGNALCVALNRRLRKYTWMEVSLAISLCVAIIVMPK 70
Qy 80 MYRS-AEHCMYGEVFCIKITSTDIMLSSASIFHLSEISIDRYAVCDPLRYAKKNILY 138
Db 71 ATEVAGGFWLFGSRCDTWAVAFDIMCSTASILHLCTISIDRYWALASPRYERRMTRL 130
Qy 139 ICVMITISVPAVAFGMIFLELNFGKAEIYYKHVHCRGCSVFFSKISGVLTPMTSF 198
Db 131 ACAMIMAAWALISLISF--VPVQLHMHKAKD---RRHSSSCDVSLNRYATLSTLSIS 184
Qy 199 YIPGSIIMLCVYRIRYILIAKEQARLISDANOKIQIGLEMKNG-----ISQSK 244
Db 185 YIPVAMITVYRIRYILIAQRIIRIST-----LERAGGOMPANSEKPASTLRSLRK 236
Qy 245 ERKAKVTGIYMGVFLICMCPFFICTVMDPLHY-----YIIPPLNDVLIWFGYLN 295
Db 237 ETKVLQTLISIMGVFCWMLPFLLNCLLPFCQPESDSNGQSPCVQGTTFENVFWMGMAN 296
Qy 296 STENPMVYAFEPYFWRKALKMML 318
Db 297 SSVPVITAFNAD-FRRAFSNTL 318

```

```

RESULT 9
A55886
dopamine receptor D1A - chicken
C:Species: Gallus gallus (chicken)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 17-Mar-1999

```

```

C:Accession: A55886
R:Demichyshyn, L.L.; Sugamori, K.S.; Lee, F.J.S.; Hamadanizadeh, S.A.; Niznik, H.B.
J. Biol. Chem. 270, 4005-4012, 1995
A:Title: The dopamine D1D receptor. Cloning and characterization of three pharmacolog
A:Reference number: A55886; MUID:95181368
A:Accession: A55886
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-450 <DEM>
A:Cross-references: GB:L36877
A:Note: authors translated the codon MET for residue 427 as Ile, and ACC for residue
C:Superfamily: vertebrate rhodopsin
C:Keywords: neurotransmitter receptor

```

```

Query Match 26.8%; Score 482; DB 2; Length 450;
Best Local Similarity 34.6%; Pred. No. 8.6e-33;
Matches 113; Conservative 56; Mismatches 110; Indels 48; Gaps 8;

Qy 31 LIIITLVGNLIVISHSFKQLHP-TNMLIHSMATVDFLGLVMPYSVRSAEHCWY 89
Db 31 LILSLTGLNTLVCAVIRFRLRSKYTNFVYISLAVSDLVAVLMPKAVSELAGFWP 90
Qy 90 FGEVECKITSTDIMLSSASIFHLSEISIDRYAVCDPLRYAKKNILYICVMIFISMSV 149
Db 91 FGS-FCNIVAVARDIMCSTASINLCVSYDRYWAISPPRYERKMPKAFIMISVAMTL 149
Qy 150 PAVAFGMIFLELNFGKAEIYYKHVHC-----RGCSVFFSKISGVLTPMTSFYIPS 203
Db 150 SVLISF--IPVQLMHKATTSFLDLNASTLQISMDNCSSLNRMVAISSLSIEFIVPA 207
Qy 204 IMCVYRIRYILIAKEQARLIS-----DANOKIQIGLEMKNGIS 241
Db 208 IMVYTRIRYIRIAKQIRKISALERAAYAKKQNTSGNRSSMDCCP-----ESNFKMS 262
Qy 242 OSKERKAVTGIYMGVFLICMCPFFICTVMDPLHY-----YIIPPLNDVLIW 291
Db 263 FKRETKVLKTLISIMGVFCWMLPFLLNCLLPFCQPESDSNGQSPCVQGTTFEIVWF 322
Qy 292 GYINSTENPMVYAFEPYFWRKALKMML 318
Db 323 GWANSSINPIIYAFNAD-FRKAESTLL 348

```

```

RESULT 10
A51660
dopamine D1B receptor - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 13-Aug-1999
C:Accession: I51660
R:Sugamori, K.S.; Demichyshyn, L.L.; Chung, M.H.; Niznik, H.B.
Proc. Natl. Acad. Sci. U.S.A. 91, 10536-10540, 1994
A:Title: D1A, D1B, and D1C dopamine receptors from Xenopus laevis.
A:Reference number: I51659; MUID:95024150
A:Accession: I51660
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-457 <SUG>
A:Cross-references: EMBL:U07864; NID:9559761; PIDN:AAA50829.1; PID:9559762
C:Superfamily: vertebrate rhodopsin
C:Keywords: neurotransmitter receptor

```

```

Query Match 26.7%; Score 481; DB 2; Length 457;
Best Local Similarity 35.5%; Pred. No. 1.1e-32;
Matches 113; Conservative 54; Mismatches 121; Indels 30; Gaps 7;

Qy 27 SLMLVLIITLVGNLIVISHSFKQLHP-TNMLIHSMATVDFLGLVMPYSVRSAB 85
Db 46 SLMLLIIFMTLEFNIIIVCAVIRFRLRSKYTNFVYISLAVSDLVAVLMPKAVAEVA 105
Qy 86 HCWYFGEVECKITSTDIMLSSASIFHLSEISIDRYAVCDPLRYAKKNILYICVMIF 145

```

Db 106 GHWPEG-APCDIIVAFDLMCTASINLCVISVDRYWAISSPRRYEKMTQVALLMIST 164
 QY 146 SMSVPAVFAFGMIFLELNFGKAEIYYKHHVHRCGCGSVFFSKISGVLTFMTSPYIPGSI 204
 Db 165 AMLAVLSLIF--IPVQLSMKSETEDHLNSHSTGNCDDSLNRYAIASSSLISFYIPVAI 222
 QY 205 MLCVYRIYLIARQEARLISDANKIQLIGLEMKNG-----ISQSKERRAKV 250
 Db 223 MIYTYRIYRIAQIQIRIKRISTLERAEHAQSCRSNRVDSCSRHHQSTLRFISIKKETVKL 282
 QY 251 TLGIVMGVFLICMPFFICTVMDPFLHY-----IIPPLINDVLIWFGYLNSTFNP 300
 Db 283 TLIIINGVFCVCMVLPFLLNCWVPFCDRSPGHQAGLPCVSETTFLEIVMGANSSLNP 342
 QY 301 MYVAFYPMFRKALKMML 318
 Db 343 IITAFNAD-PRKVFSSLL 359

RESULT 11

151659
 dopamine D1A receptor - African clawed frog (fragment)
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 13-Aug-1999
 C:Accession: I51659
 R: Sugamori, K.S.; Demchyshyn, L.L.; Chung, M.H.; Niznik, H.B.
 Proc. Natl. Acad. Sci. U.S.A. 91, 10536-10540, 1994
 A:Title: D1A, D1B, and D1C dopamine receptors from Xenopus laevis.
 A:Reference number: I51659; MUID:95024150
 A:Accession: I51659
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-451 <SUG>
 A:Cross-references: EMBL:U07863; NID:g559759; PIDN:AAA50828.1; PID:g559760
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: neurotransmitter receptor

Query Match 26.7%; Score 480; DB 2; Length 451;
 Best Local Similarity 34.5%; Pred. No. 1.3e-32;

Matches 111; Conservative 56; Mismatches 117; Indels 38; Gaps 7;

QY 31 LIITLVGNLIVISISHFKQLHTP-TNMLIHSMATVDFLLGLVMPYSMVRSAEHCWY 89
 Db 31 VLITSLTGLMTLCAAVIRRHLSKVTNFEVSLAVSDLLVAVLMPMKVAELIAGWMP 90
 QY 90 FGEVFCIKIHSTDMLSASIFHLSTIDRYAVCDPLRKAKAMNLLVCVMTFISMSV 149
 Db 91 FG-TFCNIMWAFDLMCTASTILNCVSDRYMAISSPFERYERKMPKVAEIMIGVMTL 149
 QY 150 PAVFAFMIFLELNFGKAEIYY-----KHVHRCGCGSVFFSKISGVLTFMTSPYIPGS 203
 Db 150 SVLSIF--IPVQLSMKSETEDHLNSHSTGNCDDSLNRYAIASSSLISFYIPVA 207
 QY 204 IMLCVYRIYLIARQEARLISDANKIQLIGLEMKNGISQ-----SKER 246
 Db 208 IMIYTYRIYRIARQIRISALERAHVAKNQNSTSNRSLDQOPESSLKTSFKRET 267
 QY 247 KAVTTLGIVMGVFLICMPFFICTVMDPFLH-----YIIPPLINDVLIWFGYLN 296
 Db 268 KVLTLGVIMVFCVCMVLPFLLNCVIFCDPSLTSTGTEPFCISSTTFDVFVWGMS 327
 QY 297 TFNPMVAFYPMFRKALKMML 318
 Db 328 SLNPIITAFNAD-PRKAFSSNLL 348

RESULT 12

serotonin receptor 1D alpha - rabbit
 S68423
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999

C:Accession: S68423
 R:Harwood, G.; Lockyer, M.; Giles, H.; Fairweather, N.,
 FEBS Lett. 377, 73-76, 1995
 A:Title: Cloning and characterisation of the rabbit 5-HT(1D-alpha) and 5-HT(1D-beta)
 A:Reference number: S68422; MUID:96130324
 A:Accession: S68423
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-377 <HAR>
 A:Cross-references: EMBL:Z50162; NID:g1004281; PIDN:CAA90530.1; PID:g1004282
 C:Superfamily: octopamine receptor type I
 C:Keywords: neurotransmitter receptor

Query Match 26.4%; Score 475; DB 2; Length 377;
 Best Local Similarity 32.0%; Pred. No. 2.8e-32;

Matches 112; Conservative 62; Mismatches 120; Indels 56; Gaps 7;

QY 21 VRASLVSIMVLIITLVGNLIVISISHFKQLHTPPTNMLIHSMATVDFLLGLVMPYSM 80
 Db 37 LKISLAVVLSIIVATVLSMTFVLTTLTRKLTPTANVYLGSLATYDVLVSLVMPISI 96
 QY 81 VRSAEHGWYGEVNCIKHTSDIMLSASTFHLSTIDRYAVCDPLRKAKAMNLLVIC 140
 Db 97 AYTTHTWNGQVLCDIWVSSDICTTASTILNCVLDRIWALTDLVLSKRTAGHAA 156
 QY 141 VMPIFSVPAVFAFGMIFLELNFGKAEIYYKHHVHRCGCGSVFFSKIS-GVLTFMTSPY 199
 Db 157 AMIIVNVAISICISIPPLFMR-QAKAHEV-----SDCLVNTSDISITVITGCAFY 207
 QY 200 IPGSIMLCVYRIYLIARQEQ-----ARLISDA----- 226
 Db 208 IPSVLLIVLGRITYMAARNRILNPSSLYGKRFYTAHLITGSGSSLSLSPSGEGSHS 267
 QY 227 -----NOKIQLGLEMKNGISQSKERRAVTTLGIVMGVFLICMPFFICTVMDPFL 276
 Db 268 AGSDLPFNVPARKLADSVLERKRISAAREKATVTLGILGAFGCMVPEFVASLVPIC 327
 QY 277 H-YIIPPLINDVLIWFGYLNSTFNPVAFYPMFRKALKMML-FGKIE 323
 Db 328 RDSQWMPRGLDFDFTWLGYSNLINPIITYVFNDFDQAFORVIRHFKAF 377

RESULT 13

serotonin receptor 1D - human
 A53279
 N:Altenate names: 5-hydroxytryptamine receptor 1D (5-HT1D); 5HT-1D alpha receptor
 C:Species: Homo sapiens (man)
 C:Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 05-Nov-1999
 C:Accession: A53279; A44136
 R:Hamblin, M.W.; Metcalf, M.A.
 Mol. Pharmacol. 40, 143-148, 1991
 A:Title: Primary structure and functional characterization of a human 5-HT-1D-type se
 A:Reference number: A53279; MUID:91342595
 A:Molecule type: DNA
 A:Accession: A53279
 A:Status: 1-377 <HAM>
 A:Cross-references: GB:M81869; NID:g338023; PIDN:AAA60315.1; PID:g338024
 C:Genetics:
 A:Gene: GDB:HTRLD
 A:Cross-references: GDB:132416; OMIM:182133
 A:Map position: 1p36.3-1p34.3
 A:Introns: #status absent
 C:Superfamily: octopamine receptor type I
 C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; tran

Tue Dec 11 16:12:42 2001

us-09-633-145-2.rpr

Oy 308 PWFRRALKMMLFGKIFOKDSSRCKL 332
| : : : : :
Db 293 RDEFTGYQOLEFCCLANRNRNSHKTSL 317

Search completed: December 11, 2001, 16:09:21
job time: 98 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2001, 16:07:43 ; Search time 27.16 Seconds
(without alignments)
1820.326 Million cell updates/sec

Title: US-09-633-145-2
Perfect score: 1801
Sequence: 1 MFCHNINISCVKNMSND.....FGKIFOKDSSCKLFLLEISS 338

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_TREMBL_17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhnc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 815 | 45.3 | 306 | 4 Q9P1P5 | Q9P1P5 homo sapien |
| 2 | 813.5 | 45.2 | 343 | 4 Q9P1P4 | Q9P1P4 homo sapien |
| 3 | 677.5 | 37.6 | 337 | 4 014804 | 014804 homo sapien |
| 4 | 580.5 | 32.2 | 352 | 13 Q9YHV4 | Q9YHV4 lampetra fl |
| 5 | 545.5 | 30.3 | 328 | 13 Q9YHV8 | Q9YHV8 fuqu rubrip |
| 6 | 534 | 29.7 | 359 | 4 Q9H139 | Q9H139 homo sapien |
| 7 | 526.5 | 29.2 | 328 | 13 Q9YHV7 | Q9YHV7 fuqu rubrip |
| 8 | 511 | 28.4 | 353 | 13 Q9YHV3 | Q9YHV3 lampetra fl |
| 9 | 503 | 27.9 | 446 | 13 Q42315 | Q42315 cyprinus ca |
| 10 | 498.5 | 27.7 | 445 | 13 Q98842 | Q98842 anguilla an |
| 11 | 495 | 27.5 | 437 | 13 Q42316 | Q42316 cyprinus ca |
| 12 | 488.5 | 27.1 | 358 | 11 Q9QX37 | Q9QX37 mus musculu |
| 13 | 488.5 | 27.1 | 397 | 11 Q9D282 | Q9D282 mus musculu |
| 14 | 488.5 | 27.1 | 445 | 13 Q98841 | Q98841 anguilla an |
| 15 | 481 | 26.7 | 408 | 13 Q98844 | Q98844 anguilla an |
| 16 | 477 | 26.5 | 508 | 5 Q9VC23 | Q9VC23 drosophila |
| 17 | 476.5 | 26.5 | 458 | 13 Q98843 | Q98843 anguilla an |
| 18 | 472.5 | 26.2 | 391 | 5 Q96716 | Q96716 brachyosteo |
| 19 | 462 | 25.7 | 394 | 5 Q9NH53 | Q9NH53 aplysia cal |

| | | | | | |
|----|-------|------|-----|-----------|---------------------|
| 20 | 461 | 25.6 | 394 | 5 Q9NJS6 | Q9NJS6 aplysia kur |
| 21 | 459 | 25.5 | 377 | 6 Q9N263 | Q9N263 sus scrofa |
| 22 | 454 | 25.2 | 380 | 13 Q42317 | Q42317 cyprinus ca |
| 23 | 444.5 | 24.7 | 340 | 13 Q9YHV2 | Q9YHV2 lampetra fl |
| 24 | 443 | 24.6 | 392 | 13 Q9YHV5 | Q9YHV5 myxine glut |
| 25 | 441 | 24.5 | 405 | 11 Q9CRR2 | Q9CRR2 mus musculu |
| 26 | 435.5 | 24.2 | 365 | 6 Q9N2D8 | Q9N2D8 gorilla gor |
| 27 | 435.5 | 24.2 | 464 | 5 Q9GQ54 | Q9GQ54 aedes aegypt |
| 28 | 433 | 24.0 | 366 | 6 Q9GK12 | Q9GK12 sus scrofa |
| 29 | 433 | 24.0 | 384 | 13 Q42318 | Q42318 cyprinus ca |
| 30 | 432.5 | 24.0 | 365 | 4 Q9P204 | Q9P204 homo sapien |
| 31 | 432.5 | 24.0 | 365 | 6 Q9N2D9 | Q9N2D9 pan troglod |
| 32 | 432.5 | 24.0 | 365 | 6 Q9N2D7 | Q9N2D7 pongo pygma |
| 33 | 432.5 | 24.0 | 466 | 11 Q54913 | Q54913 mus musculu |
| 34 | 432 | 24.0 | 275 | 6 Q9BF28 | Q9BF28 erinaceus c |
| 35 | 431 | 23.9 | 275 | 11 Q99P05 | Q99P05 castor cana |
| 36 | 428.5 | 23.8 | 425 | 6 Q9MZU2 | Q9MZU2 oryctolagus |
| 37 | 428.5 | 23.8 | 429 | 6 Q9MZU3 | Q9MZU3 oryctolagus |
| 38 | 428.5 | 23.8 | 603 | 5 Q9VG57 | Q9VG57 drosophila |
| 39 | 427 | 23.7 | 273 | 6 Q9BFW9 | Q9BFW9 equus caball |
| 40 | 427 | 23.7 | 275 | 6 Q9BFX0 | Q9BFX0 okapia john |
| 41 | 426 | 23.7 | 455 | 4 Q60451 | Q60451 homo sapien |
| 42 | 425.5 | 23.6 | 429 | 4 Q13729 | Q13729 homo sapien |
| 43 | 425.5 | 23.6 | 499 | 4 Q13675 | Q13675 homo sapien |
| 44 | 425 | 23.6 | 466 | 6 Q9TSW7 | Q9TSW7 sus scrofa |
| 45 | 424.5 | 23.6 | 525 | 5 Q15970 | Q15970 dugesia jap |

ALIGNMENTS

| RESULT | 1 | PRELIMINARY: | PRT: | 306 AA. |
|--|---|--------------|------|---------|
| ID Q9P1P5 | | | | |
| AC Q9P1P5 | | | | |
| DT 01-OCT-2000 (TREMBLrel. 15, Created) | | | | |
| DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update) | | | | |
| DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update) | | | | |
| DE G-PROTEIN-COUPLED RECEPTOR 58. | | | | |
| GN GPR58. | | | | |
| OS Homo sapiens (Human). | | | | |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eumammalia; Primates; Carnivora; Hominoidea; Hominidae; Homo. | | | | |
| OC Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Hominidae; Homo. | | | | |
| OX NCBI_TaxID:9606; | | | | |
| RN [1] | | | | |
| RP SEQUENCE FROM N.A. | | | | |
| RX MEDLINE-20149852; PubMed-10684976; | | | | |
| RA Lee D.K.; Lynch K.R.; Nguyen T.; Im D.-S.; Cheng F.; Saidi V.R.; | | | | |
| RA Liu Y.; Liu I.S.C.; Heng H.H.O.; Seeman P.; George S.P.; O'Dowd B.F.; | | | | |
| RA Marchese A.; | | | | |
| RT "Cloning and characterization of additional members of the G protein- | | | | |
| RT coupled receptor family." | | | | |
| RL Biochim. Biophys. Acta 1490:311-323(2000). | | | | |
| DR EMBL; AF112460; AAF27278.1; | | | | |
| DR InterPro: IPR002106; AA_Ligase_I1. | | | | |
| DR InterPro: IPR00276; GPCR_Rhodopsn. | | | | |
| DR Pfam: PF00001; 7tm.1; 1. | | | | |
| DR PRINTS: PR00337; GPCR_Rhodopsn. | | | | |
| DR PROSITE: PS00339; AA_TRNA_LIGASE_I1_2; UNKNOWN_1. | | | | |
| DR PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; UNKNOWN_1. | | | | |
| DR PROSITE: PS0262; G_PROTEIN_RECPT_F1_2; 1. | | | | |
| DR Receptor. | | | | |
| SW SEQUENCE 306 AA; 34924 MW; 55629F613062777C CRC64. | | | | |

Query Match 45.3%; Score 815; DB 4; Length 306;
Best Local Similarity 50.0%; Pred. No. 2e-58;
Matches 158; Conservative 48; Mismatches 86; Gaps 24; Gaps 4;

OY 25 LYSLLVLLITLVNGLIVISHSKQLHPPTNMLHSMATVDPLVGLVPSYMSVSA 84
DB 1 MISFAGSIFITFISNLMATISISIFKQLHPPTNMLHSMATVDPLVGLVPSYMSV 60

| ID | Q9YHY4 | PRELIMINARY; | PRT; | 352 AA. |
|----|---|--------------|------|---------|
| AC | Q9YHY4; | | | |
| DT | 01-MAY-1999 (TREMBLrel. 10, Created) | | | |
| DT | 01-MAY-1999 (TREMBLrel. 10, Last sequence update) | | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last annotation update) | | | |
| DE | PURATIVE ODORANT RECEPTOR LOR3. | | | |
| OS | Lampetria fluvialtilis (River lamprey). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia; | | | |
| OC | Petromyzontiformes; Petromyzontilidae; Lampetria. | | | |
| OX | NCBI_Taxid=7748; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=OLFATORY; | | | |
| RA | Bergnighard A., Dyer L.; | | | |
| RT | "A novel family of ancient vertebrate odorant receptors."; | | | |
| RL | J. Neurobiol. 0:0-0(1998). | | | |
| CC | -1- SIMILARITY LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY). | | | |
| CC | -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. | | | |
| DR | EMBL, AF069546; AAC82381.1; - | | | |
| DR | InterPro; IPR000276; GPCR_Rhodpsn. | | | |
| DR | Pfam; PF00001; 7tm_1; 1. | | | |
| DR | PRINTS; PR00237; GPCRHHODPSN. | | | |
| DR | PROSITE; PS00237; G_PROTEIN_RECPT_FL1; 1. | | | |
| DR | PROSITE; PS50262; G_PROTEIN_RECPT_FL2; 1. | | | |
| QW | G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane. | | | |
| QW | SEQUENCE 352 AA; 40276 MW; 6366743DcFfa9F4E CnC64; | | | |

| | | | | |
|---------------------------|-------|-------------------|-----------|------------|
| Query Match | 32.2% | Score 580.5 | DB 13 | Length 352 |
| Best Local Similarly | 36.6% | Pred. No. 2.1e-39 | | |
| Matches 121; Conservative | 69; | Mismatches 132; | Indels 9; | Gaps 5; |

| | | | |
|----|-----|--|-----|
| Qy | 9 | NISCCNNMSDVBASYSLSLWVLIITLTVGLNVLIVSISHFKQJHPTWMLHSMATVD | 68 |
| Db | 18 | NFKCTLSLSTQPKRVNLWVFTIIITITTVGNILITVLSILFRLQJRTNVLSLSTLAD | 77 |
| Qy | 69 | FLUGCLVWPYSMVRSABEHCWTGFEVPECKIHTSTDIMLSASIFHLSTISIDRYAACDPL | 128 |
| Db | 78 | FLVGLCLMPFESVMRTAASCMWFYGOJLMCRHIMHLDYTFSTTCSIFMLCISIDRYAISDPL | 137 |
| Qy | 129 | RYKKAMNVLIVCIWEIFISMSVPAPFAEGMIF-LELNFGKAEIYKXKHHVRCGCCSVEFSK | 187 |
| Db | 138 | RYDQKRVYRILAVMLTICMG--NIIPGVSYMLKLNINGIESVVAAR-SCPDNCSVEMNV | 194 |
| Qy | 188 | ISGVLTFTSTYIFGSIMLCYVYRIITYLIAKQAPLISDANOKLQIG---LEMKNGISQS | 243 |
| Db | 195 | PFGIANSNGAIVLPMLEFIMAYRSIYVWARQARISLSGQVVASNASDLMQSKNNAM | 254 |
| Qy | 244 | K-ERKAVATLGVNGVFLICCPFFICVMDPFCHYIITPLYNLDLIFGCLNSTFENPV | 302 |
| Db | 255 | KRDHNAFTLLOIMIVVLEIYVLPPIIVAVAREPVGIMDSITVMDVYANWFYFNSRBNPIL | 314 |
| Qy | 303 | YAFIYPMWRKALKMMLGKIFQKSSRCKLF | 333 |
| Db | 315 | FASFNNRSFSAFYILMSGILIRGSGYRGTDLF | 345 |

```

RESULT      5
Q9YHV8      PRELIMINARY;      PRT;      328 AA.
ID           Q9YHV8
AC           Q9YHV8;
DT           01-MAY-1999 (TREMBLrel. 10, Created)
DT           01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT           01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE           PUTATIVE NEURONTRANSMITTER RECEPTOR.
OS           Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC           Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei
OC           Acanthopterygii; Percomorpha; Tetraodontiformes;
OC           Tetraodontidae; Takifugu.
OX           NCBI_TaxID=31033;
RN           [1]

```

RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RA Rebolu J., Gardiner K., Uze G., Lutflalla G.;
RT "Class II cytokine receptor gene cluster is fixed in amniotes."
RL Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY)
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AF083321; ARC96117.1; -;
DR InterPro: IPR000276; GPCR_Kinopsn. .
DR Pfam: PF00001; 7tm_1; 1. .
DR PRINTS: PR00237; GPCRHDOPSN.
DR PROSITE: PS00237; G_PROTEIN_REEP_F1_1; 1.
DR PROSITE: PS50262; G_PROTEIN_REEP_F1_2; 1.
KW G-Protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 328 AA: 36526 MM: 687757/BJ31398C CRC64;

| | | | | |
|---------------------------|-------|------------------|------------|------------|
| Query Match | 30.3% | Score 545.5 | DB 13 | Length 328 |
| Best Local Similarity | 35.4% | Pred No. 1.3e-36 | | |
| Matches 119; Conservative | 65; | Mismatches 127; | Indels 25; | Gaps 5 |

```

QY      4 CHNIIINSCVKNMNSNDVRASLYSLMVLIIITFTVLGVLIYVISHKOLHTPTNMLIHS 63
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      10 CFKAIUNSSCKRLMOTFSEFAVSSLLSVSLVTLVLVIISHSRQHLHTPTNLLLS 69
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      64 MATVDFLLGCILMPYSAMVSRSAHECMYFGEVEYCKIHTSTDLMSASIFHLSFISIDRYYA 123
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      70 LAVSDLVGLLVMPLEGLRIYEMCKMWLGSLMCALTPIYVSCVISAQVGNMVLISIDRYLA 129
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      124 VCDPLRYKAKNNILIVICMIFISMSVPAVFAFGMIFELNEKGAELIYKRVHCHRGCCSV 183
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      130 ICDPLLYTLYTKYTLVRRAKIALICLWACSLILYNGSILMHGLQPG-----RYNSCHGECV 183
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      184 FFESKISGLVTFMTSFIYIPGSIIMLCVYRIYIILAEQARLISDANQKQIGLEMKNGISQ- 242
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      184 VISHISGTVDLEIFVFLCECTVVMVYMRVFAAQAORVIRIQVVALAV-----NAASRV 238
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      243 -SKEKRAKTKIGIYGVGLICMCPEPFICTVNDPLHIIIPTLND-----VLIIMFGILNS 296
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      239 TKSEKKARRTGIIYTVFELLCFCPRYYPALAG-----GDTSNMLIYFVAVLSWIMINS 291
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      297 TPNPVYAFYPMFPRKALKMLFCGIFPKQDSSRCKL 332
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      292 CVNPLIYALFIPMERKSIKLLVTEKILKLYSNDIKL 327

```

```

RESULT      6
Q9H199
ID      Q9H199      PRELIMINARY;      PRT;      359 AA.
AC      Q9H199;
DT      01-MAR-2001 (TReMBLrel. 16, Created)
DT      01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT      01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE      5-HDROXYTRYPTAMINE RECEPTOR.
GN      HTR4.
OS      Homo sapiens (Human) .
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      TISSUE-BRAIN (HIPPOCAMPUS);
RA      Vilario M.T., Domenech T., Palacios J.M., Mengod G.;
RT      "Cloning and characterization of multiple human 5-HT4 receptor
RT      variants including a novel variant that lacks the alternatively
RT      spliced C-terminal exon."
RL      Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC      1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY)
CC      1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS
DR      EMBL; AJ278982; CAC22251.1; -.
DR      InterPro; IPR000276; GPCR_Rhoopsn.
DR      Pfam; PF00001; 7tm_1; 1.
DR      PRINTS; PR00237; GPCRHOODPSN.

```

DR PROSITE, PS00237; G_PROTEIN_RECIP_FL_1; 1.
DR PROSITE, PS50262; G_PROTEIN_RECIP_FL_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane
SQ SEQUENCE 359 AA; 40797 MW; C66033F518178A73 CRC64;

| QY | 230 | LOIGLEMNKNGSIQS-----KERRAVKTLGIYMGVFLIOWCPFEICTYMDPFL | 276 |
|---------------------------|--|---|--------------|
| DB | 235 | LQVHAQ00000QAPETIRIROKWTAMKREHSIAKTLGIITIGLSTMLPFTYVWVLEFP-- | 292 |
| QY | 277 | HYIIPPTLNDVLIWGYLNTSTFNPMVAFFYFPMFKALKMMIFGKIFQKDSR | 329 |
| DB | 293 | NFQNSAAVRITWTWGYISSAINPLVYTLNRPFRSAFVAIVSCKVLSSTRAR | 345 |
| RESULT | 9 | | |
| ID | 042315 | PRELIMINARY: | PRT: 446 AA. |
| AC | 042315 | | |
| DT | 01-JAN-1998 | (TREMBLrel. 05, Created) | |
| DT | 01-JAN-1998 | (TREMBLrel. 05, Last sequence update) | |
| DT | 01-JUN-2001 | (TREMBLrel. 17, Last annotation update) | |
| DE | DI43 DOPAMINE RECEPTOR. | | |
| OS | Cypripus carpio (Common carp). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; | | |
| OC | Cypriniformes; Cyprinidae; Cyprininae; Cyprinus. | | |
| OX | NCBI_TaxID:7962; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | TISSUE=RETINA; | | |
| RA | Hirano J., Archer S.N., Djamgoz M.B.A.; | | |
| RL | Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases. | | |
| CC | -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY). | | |
| CC | -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. | | |
| DR | EMBL, Y14626: CAA74970.1;-- | | |
| DR | InterPro: IPR000276; GPCR_Rhodopsn. | | |
| DR | Pfam: PF00001; 7tm_1; 1. | | |
| DR | PRINTS: PR00237; GPCRHHODOPSN. | | |
| DR | PROSITE: PS00237; G-PROTEIN_RECEP_FL1; 1. | | |
| DR | PROSITE: PS02623; G-PROTEIN_RECEP_FL2; 1. | | |
| KW | G-protein coupled receptor; Glycoprotein; Transmembrane. | | |
| SO | SEQUENCE 446 AA: 49800 MW: 417834A303627D20 CRC64; | | |
| Query Match | 27.9%: | Score 503; DB 13; Length 446; | |
| Best Local Similarity | 34.7%: | Pred. No. 5e-33; | |
| Matches 118; Conservative | 68; Mismatches 120; Indels 34; Gaps 10; | | |
| QY | 7 | IINISCVKNNMSNDPRASLYSLMVLIIITLTGNIIVIVISIFHPQLHTR-TNMLIHSMA | 65 |
| DB | 8 | VLDSSVSQKNSKRYLTGCF--LSLLIMTTLTGNLTLCVCAVTKFRLKRYINFEVLSLA | 65 |
| QY | 66 | TVDPELLGLVMPYSWVRSAEHCWYFGEVFCIKHTSTDIIMLSASIFHLISIDRYAVC | 125 |
| DB | 66 | ISDLIVAILVMPKATETIVGFMPFG-AFCDVWVAFDIMCSYASILNLCVISVDRYMAIS | 124 |
| QY | 126 | DLRLKRAKNLIIVICVMFISMSYAVVAFGMIPELNLKGADEIYKHYHVRGQ---- | 180 |
| DB | 125 | SFERERKKTTPKVAAMISVAMTLLSILASE--IPQILNMHKQAQATSYELNCTYGELEPDD | 182 |
| QY | 181 | -CSVEFSKISGVLTFMTSTSYIGSITMLCYVRYIRYLAKEQARDL-----SDANQKLO | 231 |
| DB | 183 | NDSSLNRYALSSSLISITIPVALIYTYTIRIAQKQIRISALERAASAKNRHS | 242 |
| QY | 233 | IG----LEMKNGISOS--KERRAVKTLGIYMGVFLIOWCPFEICTYMDPFLH-----Y | 278 |
| DB | 243 | MGNNGSMSESSFFKSKSFREKTKVLTKLTSIVMGVFCWMLPFLVNCMVFCFNPNESSDF | 302 |
| QY | 279 | IIPPTLNDVLIWGYLNTSTFNPMVAFFYFPMFKALKMMFL | 318 |
| DB | 303 | CISSTTFDVFWFGWANSNLPIIYAFNAD-FRKAESTLL | 341 |
| RESULT | 10 | | |
| ID | 098842 | PRELIMINARY: | PRT: 445 AA. |
| AC | 098842; | | |
| DT | 01-FEB-1997 | (TRENBLrel. 02, Created) | |

```

DF 01-FEB-1997 (TrEMBLrel. 02, last sequence update)
DT
DE 01-JUN-2001 (TrEMBLrel. 17, last annotation update)
DE DOPAMINE D1A2 RECEPTOR.
GN D1A2.
OS Anguilla anguilla (European freshwater eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguillidae; Anguilla.
OX NCBI_TaxID=7936;
RN [1]
RP
RX MEDLINE=97160583; PubMed=9006917;
RA Cardinaud B., Coudouet S., Vincent J.D., Vernier P.;
RT "Early emergence of three dopamine D1 receptor subtypes in
RT vertebrates. Molecular phylogenetic, pharmacological, and functional
RT criteria defining D1A, D1B, and D1C receptors in European eel Anguilla
RT anguilla.";
RL J. Biol. Chem. 272:2778-2787(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: U62919; AAC60069.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECP_FL_2; 1.
DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
KW G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ SEQUENCE 445 AA; 49310 MW; B4A3ADAB101C1439 CRC64;

```

| Query Match | 27.7% | Score 498.5 | DB 13 | Length 445 |
|-----------------------|--------------|-------------------|----------------|------------|
| Best Local Similarity | 34.1% | Pred. No. 1,2e-32 | | |
| Matches 118 | Conservative | 66 | Mismatches 123 | Indels 39 |
| | | | | Gaps 9 |

```

OY      8  INISCVKNNMWSNDVRASTLSL-----MVLIIITLVGNLIVIVISHPKOLHP--TNMLIH 62
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     3  LNFSTVILDSGLSETSDSVAVLVGCFSLSLIVSTLLGNLTLCVAAYKFRRI;RSKYVINEFYI 62
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY     63  SMATVDFELGLVMPYSMWVRSAEHCWYEGCEVEBCKIHTSTDIMLSASIFHLFSIDRYX 122
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     63  SLAVSDLLVALVLMVPMKAVTEVAGFEPGCS-FCNIWAVARDIMCSTASILNLCIISVDRYW 121
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY     123  AWCDFLRKAKANNILVTCWIFILMSVVPVAFAGMFLFNKGAELT-YKHVHCRC-- 179
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     122  AISSPERYRKMTPKPAFAMISVAMWLLSLISF--IPVOLLNHKKQTYSYFPHNSGYGL 179
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY     180  ---GCVFESKISGVLTFTSFYIPGCSIMLCVYRYRLTAKBOARLISDANKLOQLEPM 236
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     180  LLDNCDSILNRYYAISLSSLSIFYYIPALMIVYTRRYRAQOKLRKLSL;ERAAEASAKNR 239
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY     237  KNGISQS-----KERRAKVTGLVMGVEFLICWCPEFICTVMDPLHYIIP- 281
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     240  HNSMGSSSVETESSEPKMSFKRETKVLKLTSLVIMGVFCMPLPFLINCMPFCEQAHNP 299
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY     282  -----PILNDVLMGVLNSTPKNPVVYAFPTWEPFKALKMNL 318
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     300  GSADEPCVSTTFENVFWMFGMANSSILPLTIAFNAD-FKKAFS;SL 344
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT  11
O42316      PRELIMINARY:      PRT:      437 AA.
AC      042316:
DT      01-JAN-1998 (TREMBLrel. 05, Created)
DT      01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      D1A4 DOPAMINE RECEPTOR.
OS      Cyprinus carpio (Common carp).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC      Cypriniformes; Cyprinidae; Cyprininae; Cyprinus

```

```

OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RETINA;
RL Hirano J., Archer S.N., Djamgoz M.B.A.;
RA Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: Y14627; CA74971.1; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODOPS.
DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ
SEQUENCE 437 AA; 48707 MW; A44749475F2139CE CRC64;

Query Match          27.5%; Score 495; DB 13; Length 437;
Best Local Similarity 34.5%; Pred. No. 2,2e-32;
Matches 118; Conservative 66; Mismatches 116; Indels 42; Gaps 11;

OY 7 IINISCVKNMSNDVRLASLSMLVLIITLVGNLIVISISHKQLHP-TNMLIHSMA 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 8 VLDSSVSQNSNKRVLTCGF--LSLITLTLLGNLCAAVTKFRHLRSKVTNEFVISA 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 66 TVDFLLGLVMPYRWKSAHEHCWYRGVEVCKIHTSTDIMLSSAIFHLSPISIDRYANC 125
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 66 ISDLVAILVMPWKAASEIVGFMFG-ARCDVWVAFDICSTASILNLCVISIDRYAIS 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 126 DPLRYKAKNIILVICMIFISVPAVFAFGMIFLELNFKGAEEIYKHHVCRG----- 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 125 SPFRERKMTPRVAFIMISLAWTLSTLISF--IPVQLNHKAQ-----AELNGTGEI 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 180 --GCVFESKISGVLTMTSEYIPGSIIMLCVYRYIYLAKEQARLI-----SDANOK 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 177 PNCOSLSLRRTAIISSLSLIFPAIMLVYTRIYRIAKOIRISALERAESAANKH 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 230 LDIG----LEMGNISQS--KERKAVKTGIYMGVFLIOWCPFEICTVMDPLH----- 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 237 SSMGNNAESSESSERKMSKRREKVKLTSLVINGVFCMLDPFVLCNVPFCNPESGD 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 278 -YIIPPTLNDVLIMFGYLNSTFNPVYAFYPMFRKALMML 318
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 297 FFCISSTFPDFVFWGMASSLNPLIYA-FNAGFRKAFSILL 337
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
O9OX37 PRELIMINARY; PRT; 358 AA.
ID O9OX37;
AC O9OX37;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE HISTAMINE H2 RECEPTOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SVJ;
RL Fukushima Y., Asano T., Sugano K.;
RA Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AF019138; AAD01634.1; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODOPS.
DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.

```

```

SQ SEQUENCE 358 AA; 40284 MW; D270C2C66AB7A256 CRC64;

Query Match          27.1%; Score 488.5; DB 11; Length 358;
Best Local Similarity 34.8%; Pred. No. 6e-32;
Matches 114; Conservative 56; Mismatches 111; Indels 47; Gaps 7;

OY 30 VLIITLTIV-----GNLIYVISHKQLHTFPNMLIHSMAVDFLLGLVMPYRWKSA 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 21 ISVLTTLTIFTVAGNVVCLAVSLNRRLRSNTNCIVSLATDILLGLVMPFSAIYDL 80
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 85 EHCWYRGVEVCKIHTSTDIMLSSAIFHLSPISIDRYAVCDPLRKAKRANILVICMIF 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 81 SFKMSGQYFCNLYTSLDVMCLTASILNFMISLDRYCAVTDPLRPVLTFRVAISLV 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 145 ISKSVAVAFAGMIFELNFKGAEEIYKHHVCRG-----GCVFESKISGVLTMTSEY 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 141 FIVVISTLTLSPFLSHLGWNSRNG-----TRGNDTFKCKVQYNEVYGLDGVMTPEY 191
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 200 IFGSIMLCVYRYIYLAKEQARLISDANOKLQIGLEMKNGISQSK--ERRKAVTLCI 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 192 LPLILMCVYRYIYLAKEQAKRI-----NHISWKRAITREIKATVTLAA 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 255 VMGVFLICWCPFEICTVM-----DPLHYIIPPTLNDVLIMFGYLNSTFNPVYAFYPM 309
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 238 VMGAFVTCWPEYETAFVYRGRLDGDVAENEV---EGVILMGYANSALNPILYATLRND 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 310 FRKALMMLFGKIFOKDSRCKLFELS 337
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 294 FRMAVOOLFCKLASHNSHRTSLRLNNS 321
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
O9D282 PRELIMINARY; PRT; 397 AA.
ID O9D282;
AC O9D282;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE HISTAMINE RECEPTOR H 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=CECUM;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada S.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Stuhli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Morone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AK020259; BAB32044.1; -
DR MGD; MGI:108482; Hrn2.

```



```
Db      84  AVSEVAGYWLFGK-FCETWIAFDIMCSTASINLCIISVDRYWAIASPFRYERKMTORVA 142
Qy      140  CYMIFISMSVPAVFAFGMTFLEINFKGAEIY-----YKHVHCRCGCSVFFSKISGVLP 194
Db      143  FIMIGVAMTSLILISF--IPVOLMWHKAEEDYADDNSSNHTF--DCNASLNSYVAISSS 197
Qy      195  MTSFYIPGSIIMLCVYRYRIYLIKEQARLISDANOKLO-----IGLEMKNGISQS--K 244
Db      198  LISFYIPVYIMIGTYRIRIRIAOTQIRISSLERAVEQASHOHPSDCANENSLKTTFKK 257
Qy      245  ERKAVKTLGIVNGVELICWCPEFICTVMDPLHY-----IIPPLNDVLIIMFGYLS 296
Db      258  ETKVLKTLISIIINGVEFCWLPFEVLNCMVPFCDIGEIGDPLCVSDSTFN-IPWFQGMANS 316
Qy      297  TENPMYYAEPYPERKALKML 318
Db      317  SLNPVIYAFNAD-FRKAFFSTIL 337
```

Search completed: December 11, 2001, 16:10:27
Job time: 164 sec

PR 28-FEB-2000: 2000US-0185554.
 PR 02-MAR-2000: 2000US-0186530.
 PR 03-MAR-2000: 2000US-0186811.
 PR 09-MAR-2000: 2000US-0188114.
 PR 17-MAR-2000: 2000US-0190310.
 PR 21-MAR-2000: 2000US-0190800.
 PR 20-APR-2000: 2000US-0198568.
 PR 02-MAY-2000: 2000US-0201190.
 PR 08-MAY-2000: 2000US-0203111.
 PR 25-MAY-2000: 2000US-0207094.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Vogeli G, Wood LS, Parodi LA, Riebsch RR, Lind P, Slightrom J:
 PI Schellin KA, Kayes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM:
 XX WPI: 2001-389826/41.
 DR N-PSDB; AAH51013.
 XX
 PT New G protein-coupled receptor (ngPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 XX
 PS Claim 37; Page 92; 261pp; English.
 XX
 CC The present invention relates to novel G protein-coupled receptors
 CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
 CC sequence is one such G protein-coupled receptor. GPCRs are also known as
 CC seven transmembrane receptors and function in signal transduction. The
 CC ngPCRx coding sequences are useful for screening a human to diagnose a
 CC disorder affecting the brain or a genetic predisposition, specifically
 CC schizophrenia. ngPCRx are useful for identifying compounds useful for
 CC treating schizophrenia. Detection of ngPCRx in a sample is useful as a
 CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
 CC failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
 CC metabolic and cardiovascular diseases, proliferative disorders and
 CC hormonal disorders. Modulators of ngPCRx activity have the utility for
 CC treating neurological disorders, including schizophrenia, ADHD/ADD
 CC (attention deficit-hyperactivity disorder/attention deficit disorder),
 CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
 CC migraine and senile dementia. Additional disorders include inflammatory
 CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
 CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
 CC diseases e.g. inflammatory bowel disease.
 CC
 XX
 SQ Sequence 339 AA:
 Query Match 100.0%; Score 1801; DB 22; Length 339;
 Best Local Similarity 100.0%; Pred. No. 2.5e-200;
 Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 302 myafafywpfkrakmlfkgfkdsrcklfelss 339
 RESULT
 AAB49232
 ID AAB49232 standard; protein; 339 AA.
 AC AAB49232;
 XX
 DT 14-MAR-2001 (first entry)
 DE Human SNORF33 receptor protein.
 XX
 KW SNORF33; inflammation; arthritis; neurological disorder; infection;
 KW bone disease; respiratory disorder; asthma; cancer; cardiovascular.
 OS Homo sapiens.
 XX
 PN MO200073449-A1.
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000: 2000MO-US14654.
 XX
 PR 28-MAY-1999: 99US-0322257.
 PR 06-OCT-1999: 99US-0413433.
 XX
 PA (SYNA-) SYNAPTIC PHARM CORP.
 PI Borowsky BE, Ogozalek KL, Jones KA:
 DR WPI: 2001-025252/03.
 XX
 PT Nucleic acid encoding a mammalian (human, rat and mouse) SNORF33
 PT receptor which is useful for designing drugs for treating conditions
 PT such as a chronic and acute inflammation, arthritis, neurological
 PT disorders and microbial infections -
 XX
 PS Claim 9; Fig 6; 227pp; English.
 XX
 CC The present invention relates to a mammalian SNORF33 receptor.
 CC SNORF33 antagonists and agonists are used to treat
 CC abnormalities brought about by increased or decreased activity of the
 CC mammalian SNORF33 receptor. The receptor is useful as a tool for
 CC designing drugs for treating conditions such as a chronic and
 CC acute inflammation, arthritis, neurological disorders, microbial
 CC infections, bone diseases, respiratory disorders such as asthma,
 CC cancers, cardiovascular disorders.
 CC
 XX
 SQ Sequence 339 AA:
 Query Match 100.0%; Score 1801; DB 22; Length 339;
 Best Local Similarity 100.0%; Pred. No. 2.5e-200;
 Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 SOSKERKAVKTLGIVMGVFLICWCPFFICTYMDPFLHYIIPPTLNDVLIWEGYLNSTFNP 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 242 sqskerkavkclglvmgvflilcwcpffictymdpflhyiipptlndvliwfylnstfnp 301
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 301 MYAFEPYFWRKALKMMLFGKIFQKDSRCKLFEELSS 338
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 302 myafefywftrkalkmmlfgkifqkdsrcklfeelss 339
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 3
 ID AAG80972 standard; Protein: 296 AA.
 XX AC AAG80972:
 XX DT 28-AUG-2001 (first entry)
 XX DE Human nGPCR56 #2.
 KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disorder;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder/attention deficit disorder;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective.
 XX OS Homo sapiens.
 XX PL WC200136473-A2.
 XX PD 25-MAY-2001.
 XX PE 16-NOV-2000; 2000MO-US31581.
 XX PR 16-NOV-1999; 99US-0165838.
 XX PR 17-NOV-1999; 99US-0166071.
 XX PR 19-NOV-1999; 99US-0166678.
 XX PR 28-DEC-1999; 99US-0173396.
 XX PR 22-FEB-2000; 2000US-0184129.
 XX PR 28-FEB-2000; 2000US-0185421.
 XX PR 28-FEB-2000; 2000US-0185554.
 XX PR 02-MAR-2000; 2000US-0186530.
 XX PR 03-MAR-2000; 2000US-0186811.
 XX PR 09-MAR-2000; 2000US-0188114.
 XX PR 17-MAR-2000; 2000US-0190310.
 XX PR 21-MAR-2000; 2000US-0190800.
 XX PR 20-APR-2000; 2000US-0198568.
 XX PR 02-MAY-2000; 2000US-0201190.
 XX PR 08-MAY-2000; 2000US-0203111.
 XX PR 25-MAY-2000; 2000US-0207094.
 XX PA (PHAA) PHARMACIA & UPJOHN CO.
 XX PI Vogel I G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
 XX PI Schellin KA, Kayles PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX DR WPI: 2001-369826/41.
 XX DR N-PSDB; AAH51012.
 XX PT New G protein-coupled receptor (nGPCR-x) and its encoding
 XX PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 XX PS Claim 37; Pages 91-92; 261pp; English.
 XX CC The present invention relates to novel G protein-coupled receptors
 CC (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
 CC sequence is one such G protein-coupled receptor. GPCRs are also known as
 CC seven transmembrane receptors and function in signal transduction. The
 CC nGPCRx coding sequences are useful for screening a human to diagnose a

CC disorder affecting the brain or a genetic predisposition, specifically
 CC schizophrenia. nGPCRx are useful for identifying compounds useful for
 CC treating schizophrenia. Detection of nGPCRx in a sample is useful as a
 CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
 CC failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
 CC metabolic and cardiovascular diseases, proliferative disorders and
 CC hormonal disorders. Modulators of nGPCRx activity have the utility for
 CC treating neurological disorders, including schizophrenia, ADHD/ADD
 CC (attention deficit-hyperactivity disorder/attention deficit disorder),
 CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
 CC migraine and senile dementia. Additional disorders include inflammatory
 CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
 CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
 CC diseases e.g. Inflammatory bowel disease.
 CC XX
 SQ Sequence 296 AA;
 Query Match 84.4%; Score 1520; DB 22; Length 296;
 Best Local Similarity 100.0%; Pred. No. 8,4e-168; Jaccard's 0; Gaps 0;
 Matches 286; Conservative 0; Mismatches 0; Indels 0;
 QY 1 MPFCNNIINISGVKNMNSNDVRSLSYSLMVLITLTGNIIVIVSHPOLHPTNWL 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 2 mpfcnniinisgvknmnsndvrsalsimvllltlvgnlivysshkqjhtpntwl 61
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 IHSMATVDFLGLCVMPYSMWRSAEHCWYFGEVPECKITSTDMLSAIFPLSFISIDR 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 62 lhmatavgflilgclvmprsmwrsaechwfygevfcklhtsdmlmsasr'chlsfislidr 121
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 YYAVCDPLRYKAKNNILVLCWIFISMSVPAVFAFGMFLLELNKGAEEIYKKHVRGCG 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 122 yavacdplrykaknnilvlycwifiswspavfafgmfllelnfkaeeiyykhvchrqy 181
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 CSVPFSKISGVLTFMTSPYIPGSIIMLCVYRYRIYLAKQEARLISPAANKLOIGLEMKNCI 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 182 csvfiskisgyvlltmtcstlylpslmlcvyrytlyllakeqarlispaankloiglemkngi 241
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 SOSKERKAVKTLGIVMGVFLICWCPFFICTYMDPFLHYIIPPTLNDVLIWEGYLNSTFNP 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 242 sqskerkavkclglvmgvflilcwcpffictymdpflhyiipptlndvliwfylnstfnp 301
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 4
 ID AAB49231 standard; protein: 332 AA.
 XX AC AAB49231:
 XX DT 14-MAR-2001 (first entry)
 XX DE Rat SNORF33 receptor protein.
 KW SNORF33; inflammation; arthritis; neurological disorder; infection;
 KW bone disease; respiratory disorder; asthma; cancer; cardiovascular.
 XX OS Rattus rattus.
 XX PN WC200073449-A1.
 XX PD 07-DEC-2000.
 XX PF 26-MAY-2000; 2000MO-US14654.
 XX PR 28-MAY-1999; 99US-0322257.
 XX PR 06-OCT-1999; 99US-0413433.
 XX PA (SYNA-) SYNAPTIC PHARM CORP.
 XX PI Borowsky BE, Ogozalek KL, Jones KA;
 XX DR WPI: 2001-025252/03.
 XX CC

PT Nucleic acid encoding a mammalian (human, rat and mouse) SNOBF33 receptor which is useful for designing drugs for treating conditions PT such as chronic and acute inflammation, arthritis, neurological PT disorders and microbial infections -

PS Claim 12; Fig 4; 227pp; English.

CC The present invention relates to a mammalian SMORE3 receptor.
CC SMORE3 antagonists and agonists are used to treat
CC abnormalities brought about by increased or decreased activity of the
CC mammalian SMORE3 receptor. The receptor is useful as a tool for
CC designing drugs for treating conditions such as chronic and
CC acute inflammation, arthritis, neurological disorders, microbial
CC infections, bone diseases, respiratory disorders such as asthma,
CC cancers, cardiovascular disorders.

SQ Sequence 332 AA;

| | | | | |
|---------------------------|--------|--------------------|-----------|-------------|
| Query Match | 77.98; | Score 1403; | DB 22; | Length 332; |
| Best Local Similarity | 78.18; | Pred. No. 3.7e154; | | |
| Matches 261; Conservative | 27; | Mismatches 44; | Indels 2; | Gaps 14; |

[illegible]

| | |
|----------|-------------------------------------|
| RESULT | 5 |
| AAB49234 | |
| ID | AAB49234 standard; protein; 332 AA. |

AC AAB49234 ;

DT 14-MAR-2001 (first entry)

DE Mouse SNORF33 receptor protein.

KM SNORF33, inflammation; arthritis; neurological disorder; infection;
 KN bone disease; respiratory disorder; asthma; cancer; cardiovascular.

OS Mus musculus.

PN WO200073449-A1.

PD 07-DEC-2000.

PF 26-MAY-2000; 2000WO-US14654.

| | | |
|----|--------------|---------------|
| PR | 28-MAY-1999; | 99US-0322257. |
| PR | 06-OCT-1999; | 99US-0413433. |

PA (SYNA-) SYNAPTIC PHARM CORP.

(SYNA-) SYNAPTIC PHARM CORP.

XX Borowsky BE, Ogozalek KL, Jones KA;
PI
XX WPI; 2001-02522/03.
DR

PT Nucleic acid encoding a mammalian (human, rat and mouse) SNORF33
 PT receptor which is useful for designing drugs for treating conditions
 PT such as a chronic and acute inflammation, arthritis, neurological
 PT disorders and microbial infections -

PS Claim 15; Fig 20; 227pp; English.

CC The present invention relates to a mammalian SNORF33 receptor.
CC SNORF33 antagonists and agonists are used to treat
CC abnormalities brought about by increased or decreased activity of the
CC mammalian SNORF33 receptor. The receptor is useful as a tool for
CC designing drugs for treating conditions such as a chronic and
CC acute inflammation, arthritis, neurological disorders, microbial
CC infections, bone diseases, respiratory disorders such as asthma,
CC cancers, cardiovascular disorders.

SQ Sequence 332 AA;

| | | | | |
|-----------------------|-----------------|------------------|----------|------------|
| Query Match | 75.0% | Score 1351 | DB 22 | Length 332 |
| Best Local Similarity | 75.4% | Pred. No. 4e-148 | | |
| Matches 252 | Conservative 31 | Mismatches 49 | Indels 2 | Gaps 1 |

| | | | |
|----|-----|--|-----|
| Oy | 1 | MPFCHNITISCKYKNMSDVARSLYSILMLVITLTGLNIVIVTSISPHKOLHPITNL | 60 |
| | | 1 : : : | |
| Db | 1 | whlchaltlnshnsdswsrsvvqaslyslmslllacltvglnlvivtsishkqjhrptm | 60 |
| Oy | 61 | IHSMAIVDFLLGLCLVMPYIVMVSABCHWFGVGFQKIHSTIDMLSSAIFHLSPISIDR | 120 |
| | | 1 : : : | |
| Db | 61 | Ihsmavldfllgclmpcsmvrtwercwyfgellcvtshstldmlssaafihafstidr | 120 |
| Oy | 121 | YVAAVCDPLRKAKMNLIVICWMIIFISWSPAPAFAGMFLFNKGAEEIYVYKHVCRGG | 180 |
| | | 1 : : : | |
| Db | 121 | ycavcodplrykakinlsltlmlylvswslpavayafgmflfnlnkgyeeilyrsvqsdlyg | 180 |
| Oy | 181 | CSVFESKISGVLEPMTSPYIPGSIIMLCVYRYRIYLIAKQARLISDANOKLQIGLEKKNGI | 240 |
| | | 1 : : : | |
| Db | 181 | csvfkskvsrvlaflmfsfyipgsvmlfyuyryiyflakqarslrrtn-vqvglyeqskga | 240 |
| Oy | 241 | SOSERKAVYTLIVAGVGLVICMCFEFCITWMDPLRHIIPTPLNDVLIWFGVGNSTFPP | 300 |
| | | 1 : : : | |
| Db | 239 | pqsaketaakctligimvgvflivcewfpflctvtdpblgyivdpsindclwyfgylnsainrp | 290 |
| Oy | 301 | MVYAFYVPMFRKALKMMLFGKIFQKDDSRCKLFL | 334 |
| | | 1 : : : | |
| Db | 299 | mvyaalfypwfrlalkmvllylqklfcdssrskfll | 332 |

| | |
|----------|------------------------------------|
| RESULT | 6 |
| AAG80960 | |
| ID | AAG80960 standard; Protein; 238 AA |

AC AAG80960;

DT 28-AUG-2001 (first entry)

Human nGPCR56 #1.

KM g-protein-coupled receptor; ngPCR: seven transmembrane receptor;
KM signal transduction; schizophrenia; thyroid disorder; renal failure;
KM rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KM cardiovascular disease; proliferative disorder; hormonal disorder;
KM neurological disorder; neuronal disorder; Alzheimer's disease; cancer
KM attention deficit-hyperactivity disorder; attention deficit disorder;
KM Parkinson's disease; migraine; senile dementia; inflammatory disease;
KM rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KM neuroprotective.

OS Homo sapiens.
 XX
 XX WO200136473-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 16-NOV-2000; 2000MO-US31581.
 XX
 PR 16-NOV-1999; 99US-0165838.
 PR 17-NOV-1999; 99US-0166071.
 PR 19-NOV-1999; 99US-0166578.
 PR 28-DEC-1999; 99US-0173396.
 PR 22-FEB-2000; 2000US-0184129.
 PR 28-FEB-2000; 2000US-0185421.
 PR 28-FEB-2000; 2000US-0185554.
 PR 02-MAR-2000; 2000US-0186530.
 PR 03-MAR-2000; 2000US-0186811.
 PR 09-MAR-2000; 2000US-0188114.
 PR 17-MAR-2000; 2000US-0190310.
 PR 21-MAR-2000; 2000US-0190800.
 PR 20-APR-2000; 2000US-0198568.
 PR 02-MAY-2000; 2000US-0201190.
 PR 08-MAY-2000; 2000US-0203111.
 PR 25-MAY-2000; 2000US-0207094.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J.
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX
 DR WPI; 2001-389826/41.
 DR N-PSDB; AAH51000.
 PT New G protein-coupled receptor (ngPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 XX
 PS Claim 37; Page 86; 261pp; English.
 XX
 CC The present invention relates to novel G protein-coupled receptors
 CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
 CC sequence is one such G protein-coupled receptor. GPCRs are also known as
 CC seven transmembrane receptors and function in signal transduction. The
 CC ngPCRx coding sequences are useful for screening a human to diagnose a
 CC disorder affecting the brain or a genetic predisposition, specifically for
 CC schizophrenia. ngPCRx are useful for identifying compounds useful for
 CC treating schizophrenia. Detection of ngPCRx in a sample is useful as a
 CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
 CC failure, rheumatoid arthritis, CNS disorders, infectious diseases such as HIV-1,
 CC metabolic and cardiovascular diseases, proliferative disorders and
 CC hormonal disorders. Modulators of ngPCRx activity have the utility for
 CC treating neurological disorders, including schizophrenia, ADHD/ADD
 CC (attention deficit-hyperactivity disorder/attention deficit disorder),
 CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
 CC migraine and senile dementia. Additional disorders include inflammatory
 CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
 CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
 CC diseases e.g. inflammatory bowel disease.
 CC
 XX
 SQ Sequence 238 AA;

Query Match 67.0%; Score 1207; DB 22; Length 238;
 Best Local Similarity 100.0%; Pred. No. 1.3e-131;
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPFCNIIINISCVKNNNSNDVRSALYSLMVLIIITLVGNLIIVISVISHKOLHTPPNNWL 60
 |||||||
 DB 11 mpfcnniiniscvknnsndvrsalyslmvllitltivgnliivisvishfkqltppnwl 70
 |||||||
 QY 61 IHSNATVDFLLGCLVMPYSWVRSAEHGMVGEVCKIHTSTPDIIMLSASIFHLSFISIDR 120
 |||||||
 DB 71 lhsnatvdfllgclvmpyswvrsaeahgmvgcvckihststpdimlsasifhlsfistidr 130
 |||||||

QY 121 YYAVCDPLRYKAKNNILVICMIFISWSPAVFAFGNIFLELNKGAEEIYKKHHCRGG 180
 |||||||
 DB 131 yyavcdplrykaknnilvicmifiswspavfafgmflfelnkgaeclykhvhcrgg 190
 |||||||
 QY 181 CSVFESKISGVLTFWISFYIPGSTMLCVYRIRYLIANEQARLISDANO 228
 |||||||
 DB 191 csvfesksigvltfwtfsfyipgstmlcvyrylyliakegarllisdanq 238
 |||||||

RESULT 7
 AAB49230 standard; protein; 191 AA.
 ID AAB49230
 XX
 AC AAB49230;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Human SNORF33 receptor partial protein.
 XX
 XX SNORF33; inflammation; arthritis; neurological disorder; infection;
 XX bone disease; respiratory disorder; asthma; cancer; cardiovascular.
 XX
 OS Homo sapiens.
 XX
 PA WO200073449-A1.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000MO-US14654.
 XX
 PR 28-MAY-1999; 99US-0322257.
 PR 06-OCT-1999; 99US-0413433.
 XX
 PA (SYNA-) SYNAPTIC PHARM CORP.
 XX
 PI Borowsky BE, Ogozalek KL, Jones KA;
 XX
 DR WPI; 2001-025252/03.
 XX
 PT Nucleic acid encoding a mammalian (human, rat and mouse) SNORF33
 PT receptor which is useful for designing drugs for treating conditions
 PT such as a chronic and acute inflammation, arthritis, neurological
 PT disorders and microbial infections -
 XX
 PS Disclosure; Fig 2; 227pp; English.
 XX
 CC The present invention relates to a mammalian SNORF33 receptor.
 CC SNORF33 antagonists and agonists are used to treat
 CC abnormalities brought about by increased or decreased activity of the
 CC mammalian SNORF33 receptor. The receptor is useful as a tool for
 CC designing drugs for treating conditions such as a chronic and
 CC acute inflammation, arthritis, neurological disorders, microbial
 CC infections, bone diseases, respiratory disorders such as asthma,
 CC cancers, cardiovascular disorders.
 CC
 XX
 SQ Sequence 191 AA;

Query Match 55.4%; Score 998; DB 22; Length 191;
 Best Local Similarity 100.0%; Pred. No. 1.7e-107;
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 TVDFLLCLVMPYSWVRSAEHGMVGEVCKIHTSTPDIIMLSASIFHLSFISIDRYAVC 125
 |||||||
 DB 1 tvdfllclvmpyswvrsaeahgmvgcvckihststpdimlsasifhlsfistdryavc 60
 |||||||
 QY 126 DPLRYKAKNNILVICMIFISWSPAVFAFGNIFLELNKGAEEIYKKHHCRGCSVF 185
 |||||||
 DB 61 dplrykaknnilvicmifiswspavfafgmflfelnkgaeeilykhvhcrgscsvff 120
 |||||||
 QY 186 SKISGVLTFWISFYIPGSTMLCVYRIRYLIANEQARLISDANOQLQGLEKKNISQSK 245
 |||||||

(HNHC132) used in the method of the invention. Agonists and antagonists to HNHC132 are used to respectively enhance or inhibit HNHC132 activity. Alternatively HNHC132 polynucleotide is administered in vivo to stimulate receptor activity or a nucleic acid molecule that inhibits the expression of the nucleotide sequence encoding the receptor and/or a polypeptide that competes with the receptor for its ligand are administered to inhibit HNHC132 activity. Diseases or susceptibility to diseases related to expression or activity of HNHC132 are diagnosed by determining the extent of a mutation in the nucleotide sequence encoding HNHC132 in the genome of a subject or analysing for the extent of HNHC132 expression. Identification of compounds which bind to HNHC132 comprises contacting cells with a candidate compound and assessing the ability of the candidate compound to bind to the cells (by means of stimulation or diminution of a signal). The active agents can be used in the prevention or diminution of a signal). The active agents (e.g. bacterial, fungal, protozoan and viral infections. Agonists and antagonists can be used to treat conditions associated with HNHC132 imbalance.

Sequence 343 AA:

Query Match 45.2%; Score 813.5; DB 19; Length 343;
Best Local Similarity 46.0%; Pred. No. 9,9e-86;
Matches 159; Conservative 57; Mismatches 99; Indels 31; Gaps 5;

QY 4 CHNIN-----ISCVKNN-----MSNDVRASLYSLMLVLLITLVGNLIVYISISH 49
DB 13 cktfknkllshbqlfscpgdhvfygydwhdy-----plfgnlvlnwvsh 58
QY 50 FQQLHTPTNMLIHSMTAVFLGLCLVMPYSNVAEHCWYFGEFCKIHTSTDIMLSAS 109
DB 59 fqlhspntflflsmatctflilgfvimpyssmvsescwyfgdgfkhtsfimmlrlcs 118
QY 110 ITHLSISIDRYAVCDPLRYAKKNILVYCWIFISMSVPAVEAFGMIFLENGABE 169
DB 119 ITHLSISIDRYAVCDPLRYAKKNILVYCWIFISMSVPAVEAFGMIFLENGABE 169
QY 170 IYKKNHCHGSGSVFSGVLTSTFYTGSIIMLCVYRYLYLAKQARLISANQK 229
DB 179 -YKLLVACINFCALTNKFWGTLITFCITCPSSIMVGYLYGKIFVSKGHARVISHPEN 237
QY 230 LQIGLEMKNGISOSKERRKAVKTGLIVMGVFLICWPFICTVADPLHYIIPPTLNDVLI 289
DB 238 tkgaa--vkhlshkkrkaaktglivmgvflacwlpclfiavldpdytstplldlly 295
QY 290 WEGYLNSTFNPVYAFYFWRKALMKMLFGKIPKODSSCKLPLE 335
DB 296 wlyrlnstcnplhgfnpwfqkafkyivsgkllfshsetanlfpe 341

RESULT 12

ID AAB18765 standard; Protein; 338 AA.

AC AAB18765;

DT 22-JAN-2001 (first entry)

DE Amino acid sequence of the rat SNORF1 receptor.

KW SNORF1 receptor; inflammation; arthritis; autoimmune disease;

KW transplant rejection; infection; AIDS; pain; psychotic disorder;

XX neurological disorder.

OS Rattus norvegicus.

XX Key Location/Qualifiers

FT 23..48 /note= "putative transmembrane domain"

FT 57..78 /note= "putative transmembrane domain"

FT 101..118 /note= "putative transmembrane domain"

FT /note= "putative transmembrane domain"
FT 138..159 /note= "putative transmembrane domain"
FT Domain 188..214 /note= "putative transmembrane domain"
FT Domain 246..270 /note= "putative transmembrane domain"
FT Domain 283..308 /note= "putative transmembrane domain"

US6117990-A.

12-SEP-2000.

06-APR-1999; 99US-0286805.

06-APR-1999; 99US-0286805.

(SYNA-) SYNAPTIC PHARM CORP.

Bonini JA, Borowsky BE;

WPI; 2000-610854/58.

N-PSDB; AAA75825.

New recombinant nucleic acid molecules, with sequences identical to the human and rat SNORF1 receptor-encoding nucleic acid molecules and possessing the structural motif characteristics of a G-protein

Disclosure; Fig 4A-B; 19pp; English.

The present sequence represents a SNORF1 receptor. The SNORF1 nucleic acids may be used a probes to obtain homologous nucleic acids from other species and to detect the existence of nucleic acids having complementary sequences in samples. The nucleic acids may also be used to express the receptors they encode in transfected cells. Also, use of the receptor encoded by the SNORF1 receptor nucleic acid sequence enables the discovery of the endogenous ligand and to elucidate the role of the SNORF1 receptor. The receptor may be employed for designing drugs for treating various pathological conditions such as chronic and acute inflammation, arthritis, autoimmune diseases, transplant rejection, bacterial and fungal infections, AIDS, pain, psychotic and neurological disorders. Transfected cells with the receptor may be used to test compounds which bind to the receptor and which activate or inhibit the functional responses.

Sequence 338 AA:

Query Match 43.4%; Score 782; DB 21; Length 338;
Best Local Similarity 45.8%; Pred. No. 4,4e-82;
Matches 154; Conservative 63; Mismatches 115; Indels 4; Gaps 4;

QY 1 MPFCHNINISCVKNNMSNDVRASLYSLMLVLLITLVGNLIVYISISHFQQLHTPTNML 60

DB 1 mclcyenvnscikssyspwrailayavlglgallavfllaytallhfkqlhptnfl 60

QY 61 IHSMTAVDFLGLCLVMPYSNVAEHCWYFGEFCKIHTSTDIMLSASIFHSIDIR 120

DB 61 vasiacadtlygvtmpfsvtsvegcwyfgdgfkhtcfdgfsfaalfhccisldt 120

QY 121 YVAVCDPLRYAKKNILVYCWIFISMSVPAVEAFGMIFLENGABEITYYKHNVRG 180

DB 121 yavavcdplrypkftlsvsgvalawfsvtsfslfyganeegldelwa-lcvgg.179

QY 181 GSVEFSKISGVLTSTFYTGSIIMLCVYRYLYLAKQARLI-SANQKLOIGLEMKNG 239

DB 180 cgaplmgwvllcflf-fliprvwmvflygriflvakqgarklegsangpqasseyker 238

QY 240 ISGSKERRKAVKTGLIVMGVFLICWPFICTVADPLHYIIPPTLNDVLIIMFQYLNSTF 299

DB 239 var-rrrkaaktgliaaallvswlpyldavdaymnltpayvylwvcvynsann 297

[illegible]

| Seq | Sequence | 348 AA |
|---|--|--------|
| Query Match | 41.3% Score 743; DB 21; Length 348; | |
| Best Local Similarity | 42.5% Pred. No. 1.5e-77; | |
| Matches 141; Conservative 63; Mismatches 126; Indels 2; Gaps 2; | | |
| QY | 4 CHNININISCVNNMNSNDVRASLYSLMWLLITLTVLGNLIVIVSISHFQOLHPTNMLIHS 63 | |
| Db | 14 CYKNHVESCKLTPSPGPRSLIYAVLFGFVLAAGFNLIMLAIHFKQLHCPNFIAS 73 | |
| QY | 64 MATVDPLFLGCLVMYSESWARSAEHCWYEGEVCKIHTSDIMLSASATPHLSFISIDRYA 123 | |
| Db | 74 IACEDFLVGYVYVMPFSICVRSVESCWYFGDSYCKFHTCDDTSFCFSIHLCEISIDRYA 133 | |
| QY | 124 VCDPLRKAKNNILVICWMIFFISMSVPAVFAFGMIFLELNKGAEEIYKKHVCRGGSV 183 | |
| Db | 134 VTDPLTPKCTKEVAVSGIQLVSWFISVYFISITFTGANEAGIELVVA-ITCVGSGQA 192 | |
| QY | 184 FFESISGVLTFTMISFYIPGSIIMLCVYIYIYLAKEQALISDANOKLDIGLEMKNGISQS 243 | |
| Db | 193 PLDGNWVILLCIL-FFIPNVAMVFIYAKLIFVAKIPQAKLSTASQASQSSSYKERVAK 251 | |
| QY | 244 KERAVNTIGIVMCGVFLVCMCPFFICVYMDPEFLHIIPTLNDVYIWMGVLNSTNPWY 303 | |
| Db | 252 REIRAAKTIGLAMAAFLVSWALPYLDVADIDYMTFLTPPYVELLYVCVYNSAMPLY 311 | |
| QY | 304 AFEPYMERKALKMMLFKRIFQKDSRCKLPLE 335 | |
| Db | 312 AFFYGFYGAIKKILVSGKVLRTDSSTNLFSQ 343 | |
| RESULT 14 | | |
| AAG80970 | | |
| ID | AAG80970 standard; Protein: 345 AA. | |
| XX | | |
| AC | AAG80970; | |
| XX | | |
| DT | 28-AUG-2001 (first entry) | |
| XX | | |
| DE | Human nGPCR40 #2. | |
| XX | | |
| KW | G, protein-coupled receptor; nGPCR; seven transmembrane receptor; | |
| KW | signal transduction; schizophrenia; thyroid disorder; renal failure; | |
| KW | rheumatoid arthritis; CNS disorder; infection; metabolic disease; | |
| KW | cardiovascular disease; proliferative disorder; hormonal disease; | |
| KW | neurological disorder; neuronal disorder; Alzheimer's disease; cancer; | |
| KW | attention deficit-hyperactivity disorder/attention deficit disorder; | |
| KW | Parkinson's disease; migraine; senile dementia; inflammatory disease; | |
| KW | rheumatoid arthritis; autoimmune disorder; respiratory ailment; | |
| KW | neuroprotective. | |
| OS | Homo sapiens. | |
| XX | | |
| XX | WO200136473-A2. | |
| PN | | |
| PD | 25-MAY-2001. | |
| XX | | |
| PF | 16-NOV-2000; 2000WO-US31581. | |
| XX | | |
| PR | 16-NOV-1999; 99US-0165838. | |
| PR | 17-NOV-1999; 99US-0166071. | |
| PR | 19-NOV-1999; 99US-0166678. | |
| PR | 28-DEC-1999; 99US-0173396. | |
| PR | 22-FEB-2000; 2000US-0184129. | |
| PR | 28-FEB-2000; 2000US-0185421. | |
| PR | 28-FEB-2000; 2000US-0185554. | |
| PR | 02-MAR-2000; 2000US-0186530. | |
| PR | 03-MAR-2000; 2000US-0186811. | |
| PR | 09-MAR-2000; 2000US-0188114. | |
| PR | 17-MAR-2000; 2000US-0190310. | |
| PR | 21-MAR-2000; 2000US-0190800. | |
| PR | 20-APR-2000; 2000US-0198568. | |

Tue Dec 11 16:12:40 2001

us-09-633-145-2.rag

Page 11

Search completed: December 11, 2001, 16:09:52
Job time: 129 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2001, 16:08:44 : Search time 13.2 seconds
(without alignments)
938.842 Million cell updates/sec

Title: US-09-633-145-2

Perfect score: 1801
Sequence: 1 MPFCHNINISCVKNMNSND.....FGKIFOKDSSCKLEFLSS 338

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 542.5 | 30.1 | 388 | 5H4_CAVPO | O70528 cavia porce |
| 2 | 539 | 29.9 | 388 | 5H4_MOUSE | P97288 mus musculu |
| 3 | 534 | 29.7 | 388 | 5H4_HUMAN | Q13633 homo sapien |
| 4 | 529 | 29.4 | 406 | 5H4_RAT | O62758 rattus norv |
| 5 | 510.5 | 28.3 | 459 | 1 DDDR_FUGRU | P33452 figu rubrip |
| 6 | 497.5 | 27.6 | 465 | 1 DDDR_XENLA | P42291 xenopus lae |
| 7 | 493.5 | 27.4 | 363 | 1 DDDR_CARAU | P35406 carassius a |
| 8 | 492.5 | 27.3 | 358 | 1 HH2R_MOUSE | P97292 mus musculu |
| 9 | 486 | 26.7 | 358 | 1 HH2R_RAT | P25102 rattus norv |
| 10 | 481 | 26.7 | 457 | 1 DDDR_XENLA | P42290 xenopus lae |
| 11 | 480 | 26.7 | 451 | 1 DDDR_XENLA | P42289 xenopus lae |
| 12 | 475 | 26.4 | 377 | 1 SH1D_RABIT | P91945 oryctolagus |
| 13 | 474 | 26.3 | 377 | 1 SH1D_HUMAN | P28221 homo sapien |
| 14 | 473 | 26.3 | 374 | 1 SH1D_RAT | P28565 rattus norv |
| 15 | 472.5 | 26.2 | 359 | 1 HH2R_HUMAN | P25021 homo sapien |
| 16 | 471 | 26.2 | 446 | 1 DADR_DIDMA | P42288 didelphis m |
| 17 | 469 | 26.0 | 374 | 1 SH1D_MOUSE | O61224 mus musculu |
| 18 | 467 | 25.9 | 446 | 1 DADR_PIG | P50130 sus scrofa |
| 19 | 466.5 | 25.9 | 377 | 1 SH1D_CANFA | P11614 canis fami1 |
| 20 | 466.5 | 25.9 | 386 | 1 DDDR_ORMO | P47800 oreochromis |
| 21 | 462.5 | 25.7 | 463 | 1 D5DR_FUGRU | P34554 figu rubrip |
| 22 | 461.5 | 25.6 | 376 | 1 SH1D_CAVPO | O60484 cavia porce |
| 23 | 461 | 25.6 | 483 | 1 B1AR_MELGA | P07700 melagris g |
| 24 | 459.5 | 25.5 | 359 | 1 HH2R_CAVPO | P47747 cavia porce |
| 25 | 458 | 25.4 | 446 | 1 DADR_HUMAN | P21728 homo sapien |
| 26 | 457 | 25.4 | 446 | 1 DADR_MACMU | O77680 macaca mula |
| 27 | 454.5 | 25.2 | 477 | 1 DDDR_HUMAN | P21918 homo sapien |
| 28 | 453 | 25.2 | 385 | 1 B1AR_XENLA | O42574 xenopus lae |
| 29 | 452.5 | 25.1 | 379 | 1 SH1D_FUGRU | P79748 figu rubrip |
| 30 | 452.5 | 25.1 | 418 | 1 B2AR_BOVIN | O28044 bos taurus |
| 31 | 452 | 25.1 | 418 | 1 B2AR_MESAU | P04274 mesocricetu |
| 32 | 451 | 25.0 | 418 | 1 B2AR_RAT | P10608 rattus norv |
| 33 | 450.5 | 25.0 | 359 | 1 HH2R_CANFA | P17124 canis fami1 |

| | | | | | | |
|----|-------|------|-----|---|------------|--------------------|
| 34 | 450 | 25.0 | 475 | 1 | DBDR_RAT | P25115 rattus norv |
| 35 | 449 | 24.9 | 418 | 1 | B2AR_MOUSE | O18762 mus musculu |
| 36 | 448 | 24.9 | 467 | 1 | B1AR_BOVIN | O31962 bos taurus |
| 37 | 446.5 | 24.8 | 415 | 1 | B2AR_MACMU | O28509 macaca mula |
| 38 | 446.5 | 24.8 | 446 | 1 | DADR_RAT | P18901 rattus norv |
| 39 | 446 | 24.8 | 466 | 1 | B1AR_MOUSE | P34971 mus musculu |
| 40 | 446 | 24.8 | 466 | 1 | B1AR_RAT | P18090 rattus norv |
| 41 | 445 | 24.7 | 467 | 1 | B1AR_SHEEP | O28927 ovis aries |
| 42 | 444.5 | 24.7 | 418 | 1 | B2AR_FELCA | O28515 felis silve |
| 43 | 444.5 | 24.7 | 468 | 1 | B1AR_PIG | O28998 sus scrofa |
| 44 | 443.5 | 24.6 | 415 | 1 | B2AR_CANFA | P54833 canis fami1 |
| 45 | 443.5 | 24.6 | 474 | 1 | B1AR_FELCA | O91616 felis silve |

ALIGNMENTS

| RESULT | ID | 5H4_CAVPO | STANDARD: | PRT: | 388 AA. |
|--------|--|-----------|-----------|------|---------|
| AC | O70528: | | | | |
| DT | 30-MAY-2000 (Rel. 39, Created) | | | | |
| DT | 30-MAY-2000 (Rel. 39, Last sequence update) | | | | |
| DT | 30-MAY-2000 (Rel. 39, Last annotation update) | | | | |
| DE | 5-HYDROXYTRYPTAMINE 4 RECEPTOR (5-HT4). | | | | |
| DE | 5-HYDROXYTRYPTAMINE 4 RECEPTOR (5-HT4). | | | | |
| GN | HT4. | | | | |
| OS | Cavia porcellus (Guinea pig). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Rodentia; Hysticomnathi; Caviidae; Cavia. | | | | |
| OX | NCBI TaxID=10141; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RA | Van den Wyngaert I., Gommeren W., Jurzak M., Verhaesselt P., Gordon R., | | | | |
| RA | Leyssen J., Luyten W., Bender E.; | | | | |
| RT | "Cloning and expression of 5-HT4 receptor species and splice | | | | |
| RT | variants."; | | | | |
| RL | Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases. | | | | |
| CC | -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR | | | | |
| CC | 5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC AMINE THAT FUNCTION | | | | |
| CC | AS A NEUOTRANSMITTER, A HORMONE, AND A MITOCH. ACTIVITY OF | | | | |
| CC | THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT S-FUNCTIONS ADENYLATE | | | | |
| CC | CYCLASE (BY SIMILARITY). | | | | |
| CC | -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. | | | | |
| CC | -1- ALTERNATIVE PRODUCTS: A NUMBER OF FORMS ARE PRODUCED BY | | | | |
| CC | ALTERNATIVE SPLICING. | | | | |
| CC | -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. | | | | |
| CC | THIS SWISS-PROT entry is copyright. It is produced through a collaboration | | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL Outstation - | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | |
| CC | entities requires a license agreement (see http://www.isb-sib.ch/announce/ | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | |
| CC | ----- | | | | |
| DR | EMBL: Y13585; CAA73912.1; | | | | |
| DR | InterPro: IPR000276; GPCR_Rhodopsn. | | | | |
| DR | Pfam: PF00001; 7tm.1; 1. | | | | |
| DR | PRINTS: PR00237; GPCRHHODOPSN. | | | | |
| DR | PRINTS: PR01059; 5HT4RECEPTR. | | | | |
| DR | PRINTS: PR01103; ADRENERGICR. | | | | |
| DR | PROSITE: PS00237; G-PROTEIN_RECEP_F1.1; 1. | | | | |
| DR | PROSITE: PS50262; G-PROTEIN_RECEP_F1.2; 1. | | | | |
| KW | G-protein coupled receptor; transmembrane; Glycoprotein; | | | | |
| KW | Multigene family; Lipoprotein; Palmitate; Alternative splicing. | | | | |
| FT | DOMAIN 1 | | | | |
| FT | TRANSSEM 20 | | | | |
| FT | TRANSSEM 41 | | | | |
| FT | DOMAIN 59 | | | | |
| FT | TRANSSEM 79 | | | | |
| FT | DOMAIN 80 | | | | |
| FT | TRANSSEM 94 | | | | |
| FT | TRANSSEM 117 | | | | |
| FT | DOMAIN 137 | | | | |
| FT | TRANSSEM 158 | | | | |


```

FT LIPID 329 329 PALMITATE (BY SIMILARITY).
FT VANSPLC 169 169 L-> LERSLNGLGODPHA (IN ISOFORM 5-
FT VANSPLC 359 388 HT4(E)).
FT VANSPLC 359 388 RDVCEGGWESQCHPPTSPLVAAPSDT -> SGCSPVS
FT VANSPLC 359 388 SELEFCNRPVPV (IN ISOFORM 5-HT4(E)).
FT VANSPLC 359 388 RDVCEGGWESQCHPPTSPLVAAPSDT -> SSGFTD
FT VANSPLC 359 388 RNFGRIRRLTKPS (IN ISOFORM 5-HT4(D)).
FT VANSPLC 360 388 DAVECGGWESQCHPPTSPLVAAPSDT -> F (IN
FT VANSPLC 360 388 ISOFORM 5-HT4(C)).
FT VANSPLC 360 388 DAVECGGWESQCHPPTSPLVAAPSDT -> YTVLRGH
FT VANSPLC 360 388 HOELKRLPHNDPESLESCF (IN ISOFORM 5-
FT VANSPLC 360 388 HT4(A)).
SQ SEQUENCE 388 AA; 43761 MW; 7FCFEC60E7BDF560 CRC64;

Query Match 29.7%; Score 534; DB 1; Length 388;
Best Local Similarity 37.4%; Pred. No. 2.1e-27;
Matches 116; Conservative 53; Mismatches 121; Indels 20; Gaps 5;

OY 25 LYSIAVLITLTGVLNGLIYVSIHSHKOLH-TPTNMLHSMATVDFLLGCLVWPYMSVRS 83
DB 22 LITFISTYLMAIIGLWLVAVAVCNDROLRKIKTNFTVSLAFADLLVSLVWPFQAIEL 81
OY 84 AEHCWTFGEVFCIKINTSDIMLSASIFHLSEISIDRYAV-CDPLRYKAKNIIYICW 142
DB 82 VQDIWIYGEVFLVRSIDVLITFTASIFHLCCSIDRYAICQPLVYNNKMTPLRIALM 141
OY 143 ITISVSVPAVPAFGMIFELNKGAEIYKHYHCRGG---CSVFSSISGVLTPTMTSF 198
DB 142 LGGCWVITPFIPLFMQNMNIGIIDLEKRFNONSSTYCVFVWNPYATITCSVAF 201
OY 199 YIPGSIIMCVYRIYLIAKEQARLISDANQKLOIGLEMNKGISQS-----KERKA 248
DB 202 YIPELLMVAIYRIYTAKEHNOI---QMLORAGASSESPQSDQSHTHMRPETYA 257
OY 249 VKTGLVGMVFLICWCPFEICTMDPEFLHYIIPPLNDVILMGVINSTFNWVAFYFP 308
DB 258 AKTLCIMGCFICWAPFVTVNIDPFIYVPGQWVAFMLGVIYNSGILPFLYAFUNK 317
OY 309 WFRKALKMML 318
DB 318 SFRRAFLLIT 327

RESULT 4
SH4_RAT STANDARD: PRT; 406 AA.
AC 062758; 062757; 063006; 089034;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 5-HYDROXYTRYPTAMINE 4 RECEPTOR (5-HT-4) (SEROTONIN RECEPTOR) (5-HT4).
CN HTR4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RA MEDLINE=95317299; PubMed=7796807;
RA Gerald C., Adam N., Kao H.T., Olsen M.A., Laz T.M., Schechter L.E.,
RA Bard J.A., Vaysse P., Hartig P.R., Branchek T.A., Weisshank R.L.;
RA "The 5-HT4 receptor: molecular cloning and pharmacological
RT characterization of two splice variants.";
RL EMO J. 14:2806-2815(1995).
RN [2]
RP SEQUENCE OF 165-259 FROM N.A.
RC TISSUE-Brain;
RA MEDLINE=95385798; PubMed=7656980;
RA Ullmer C., Schmuck K., Kalkman H.O., Lubbert H.;
RA "Expression of serotonin receptor mRNAs in blood vessels.";
RL FEBS Lett. 370:215-221(1995).

```

```

RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 5-HT4(E)).
RC TISSUE-Brain;
RA MEDLINE=95238795; PubMed=10220570;
RA Claessen S., Sebben M., Becamel C., Bockaert J., Dumuis A.;
RA "Novel brain-specific 5-HT4 receptor splice variants show marked
RT constitutive activity: role of the C-terminal intracellular domain.";
RL Mol. Pharmacol. 55:910-920(1999).
CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR
CC 5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTION
CC AS A NEUROTROPHIC AGENT, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT STIMULATES ADENYLATE
CC CYCLASE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS: 5-HT4L (SHOWN HERE),
CC 5-HT4S AND 5-HT4(E) ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: IN BRAIN, THE 5-HT4S TRANSCRIPTS ARE
CC RESTRICTED TO THE STRIATUM, BUT THE 5-HT4L TRANSCRIPTS ARE
CC EXPRESSED THROUGHOUT THE BRAIN, EXCEPT IN THE CEREBELLUM. IN
CC PERIPHERAL TISSUES, DIFFERENTIAL EXPRESSION IS ALSO OBSERVED IN
CC THE ATRIUM OF THE HEART WHERE ONLY THE 5-HT4S ISOFORM IS
CC DETECTABLE.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
CC EMBL: 248153; CAA8170.1; -
CC EMBL: AJ011370; CAA09599.1; -
CC DR GCRDB: GCR_1447; -
CC DR GCRDB: GCR_1448; -
CC DR GCRDB: GCR_1480; -
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PRINTS: PR01059; 5HT4RECEPTR.
CC DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1.1; 1.
CC DR PROSITE: PS50262; G_PROTEIN_RECPEP_F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Palmitate.
CC MultiGene family; Alternative splicing; Lipoprotein; Palmitate.
CC -----
CC EMBL: U20907; AAC52232.1; -
CC EMBL: U20906; AAC52232.1; -
```


Query Match 29.4%; Score 529; DB 1; Length 406;
 Best Local Similarity 35.4%; Pred. No. 4, 5e-27;
 Matches 111; Conservative 61; Mismatches 114; Indels 28; Gaps 6;

25 LYSIMVLIITLTVGNLIVIVSISHEFKQLH-TPTNMLIHSNATVDFLLGCLVMPYSNVR 83
 22 LITFFPAVLIIMAILGNLVAVACRDRQLRKIKITNFIYSIAFLADLVLSVNAFGAIEL 81
 84 AEHCWGEVEFECKIHTSTDIIMLSASIFHLSTISIDRYAV-CDPLRYKAKMNLIVICVM 142
 82 VQDIWEFGEHFCVRSITLDVLLTASIFHLCSIDRYAACCPPLVYRNKMTPLRLTALM 141
 143 IFTSMSPVAPFACGMFLLEINFGAEIY----YKHVHCRCGCVSFPSKISGVLTMTSF 198
 142 LGCGWVIVPMFISFLPIQGNMNIIVDIVIEKRKNHNSNTEFCVFNKPYAITCSYVAF 201
 199 YIGSIMLCVYRIRYLAKQARLISDANKLQIGLEMKNGISQK----- 244
 202 YIPPLMVLIYRIRYVAKHA-----QOIQM-IQPAQATSSRQTDADQSTHNMRT 253

245 ERRAKVTIGVGVFLICMCFEFCIYMDPLATIPPLNDVLIIMFGYLNSTFNPMVYA 304
 254 ETRAKATLCVIMGCFCCMAPEFYTNIVDPIDTYVEKWTATLWLGITNSGLNPLIYA 313
 305 FEPYWERKALKMML 318
 314 FLNKSFRRAFLITL 327

RESULT 5
 DDIR_FUGRU STANDARD; PRT; 459 AA.
 AC P53452;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE D(1)-LIKE DOPAMINE RECEPTOR.
 GN D14.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Takifugu.
 OX NCBI_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95309911; PubMed=7789977;
 RA Machae A.D., Brenner S.;
 RT "Analysis of the dopamine receptor family in the compact genome of the puffer fish Fugu rubripes."
 RL Genomics 25:436-446(1995).
 CC -|- FUNCTION: RECEPTOR FOR DOPAMINE.
 CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X80174; CA56455.1; -
 DR HSSP: P29274; 1MMH.
 DR GCRDB: GCR_1108; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PRINTS: PR00242; DOPAMINER.
 DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 MultiGene Family.
 FT DOMAIN 1 23
 FT TRANSSEM 24 49
 FT DOMAIN 50 60
 FT TRANSSEM 61 87
 FT DOMAIN 88 96
 FT TRANSSEM 97 119
 FT DOMAIN 120 138
 FT TRANSSEM 139 164
 FT DOMAIN 165 191
 FT TRANSSEM 192 216
 FT DOMAIN 217 269
 FT TRANSSEM 270 297
 FT DOMAIN 298 311
 FT TRANSSEM 312 333
 FT DOMAIN 334 459
 FT CARBOHYD 4 4
 FT DISULFID 96 187
 SQ SEQUENCE 459 AA; 51080 MW; B69857A3A4E10B CRC64;

Query Match 28.3%; Score 510.5; DB 1; Length 459;
 Best Local Similarity 37.0%; Pred. No. 7, 2e-26;
 Matches 118; Conservative 54; Mismatches 110; Indels 37; Gaps 8;

31 LIITLTVGNLIVIVSISHEFKQLHTP-TNWLIHSNATVDFLLGCLVMPYSNVRSAEHCWY 89
 32 LITFTLGLNLTVCVAVYKFRHLRSKVTNFFVLSAISDLVALVIMWKATATEMGWMP 91
 90 FGEVFECKIHTSTDIIMLSASIFHLSTISIDRYAVCDPLRYKAKMNLIVICVMFISNSV 149
 92 FGE-FCNIMVAFDIMCSASTILNCVTSIDRYWMAISSPEFRERKMTPEVACIMTSVWTL 150
 150 PAVFAFGMIFLELNPKGAKEELYYKHHVHORG-----GCVSFPSKISGVLTMTSFYI 200
 151 SVLSF--IPQLWVHKAQTASY--VELNGTYAGDLPDNCDSLNRYAISSSLISFYI 206
 201 PGSIMLCVYRIRYLAKQARLI-----SDANKLQIGLEMKNGISQSKER 246
 207 PVALIMVYTHYTHYRAKQIRISALERAESAQRNHSMSGNSLSMECECFKMSFKRET 266
 247 KAVTGLGVGVFLICMCFEFCIYMDPLATIPPLNDVLIIMFGYLNSTFN 299
 267 KVLKTLVIMGVFVCCWPEFLNCMPFCEADDTTDPCCISSTTFDVFVFGWANSLSN 326

300 PMVYAFEPYWERKALKMML 318
 327 PITYAFNAD-FRKAFSTILL 344

RESULT 6
 DCDR_XENLA STANDARD; PRT; 465 AA.
 AC P42291;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE D(1C) DOPAMINE RECEPTOR.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95024150; PubMed=7937989;
 RA Sugamori K.S., Demchishyn L.L., Chung M., Niznik H.B.;
 RT "D1B, D1B, and D1C dopamine receptors from Xenopus laevis."
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10536-10540(1994).
 CC -|- FUNCTION: THIS IS ONE OF THE FIVE TYPES (D1 TO D5) OF RECEPTORS FOR DOPAMINE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.

```

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: BRAIN AND KIDNEY.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
-----
DR EMBL: U07865; AAA50830.1; -.
DR HSSP: P29274; 1MMH.
DR GCRDB: GCR_1236; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PRINTS: PR00242; DOPAMINER.
DR PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE: PS00262; G-PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Lipoprotein; Palmitate.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 54 1 (POTENTIAL).
FT DOMAIN 55 65 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 66 92 2 (POTENTIAL).
FT DOMAIN 93 101 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 102 124 3 (POTENTIAL).
FT DOMAIN 125 143 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 144 168 4 (POTENTIAL).
FT DOMAIN 169 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 219 5 (POTENTIAL).
FT DOMAIN 220 264 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 265 291 6 (POTENTIAL).
FT DOMAIN 292 309 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 310 334 7 (POTENTIAL).
FT DOMAIN 335 465 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 3 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 8 8 BY SIMILARITY.
FT LIPID 101 187 PALMITATE (BY SIMILARITY).
FT LIPID 344 344
SQ SEQUENCE 465 AA; 52640 MM; F4IDF85AF0D2P869 CRC64;

```

```

Query Match 27.6%; Score 497.5; DB 1; Length 465;
Best Local Similarity 34.5%; Pred. No. 4.8e-25;
Matches 118; Conservative 61; Mismatches 130; Indels 33; Gaps 8;

```

```

QY 6 NIINISCVKNNMSND-----VRASLYSLMVLIIITLVGNLIVIVSISFKQLHNP 56
DB 5 STEFNTV--NWVHADLDVGNDSLSLRALTLGLLSLILSLTLGNLIVCLAVIKFRLLRSK 62
QY 57 -TNWLIHSNATVDFLLGLVMPYSNVRSAEHGWCYFGEVCKIHTSTDMLSSASIFHLSE 115
DB 63 VTNFVVISLAVSDLEFVALLVMPWKAVTEVAGFWFGD-FCDTWVAFDIDICSTASILNLCI 121
QY 116 ISIDRYVAVCPDLRYKAKMNIIVICMIFISWSVPAVRAFGIPELANKGAEIYYKAV 175
DB 122 ISLDRTYMAISPFERYERKRTQVAFIMIGVAMTSLISLIFIVOLSWHNSHRADELNGV 181
QY 176 HCRGGSVFEFSKISGLVFMFTSFIPIGSIIMLCVYRIYLIAKEQARLISDAN-----OK 229
DB 182 NHTENCDSLNRTVIAISSLSIFYPVIMITRYIRIRIAQTQIRRISSLEBRAVHAOR 241
QY 230 LQIGLEKNGISQ--KERRAKVTLGIWGVFLICWCPEFFICTVADPELHYIIP----- 281
DB 242 CSSRLSNENSLKTSERKETKVKLTLSIMGVFCMLPEFLVNCMIIPCHMMLPGQNEDE 301
QY 282 -----PLINDVLIWFGYLNSTFNPNVYAFYFVFRKALKMML 318
DB 302 PCVSETTFNLFVWFGMANSSLNPVIAFNAD-FRKAFTTIL 342

```

```

RESULT 7
DIDR_CARAU STANDARD: PRT; 363 AA.
AC P35406;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE D(1) DOPAMINE RECEPTOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=retina;
RX MEDLINE=94088471; PubMed=8264547;
RA Frail D.E., Manelli A.M., Witte D.G., Lin C.W., Steffey M.E.,
RA Mackenzie R.G.;
RT "Cloning and characterization of a truncated dopamine D1 receptor
RT from goldfish retina: stimulation of cyclic AMP production and
RT calcium mobilization."
RL Mol. Pharmacol. 44:1113-1118(1993).
CC -1- FUNCTION: THIS IS ONE OF THE FIVE TYPES (D1 TO D5) OF RECEPTORS
CC FOR DOPAMINE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G
CC PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE. COULD BE INVOLVED IN
CC GROWTH HORMONE RELEASE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: RETINA.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
-----
DR EMBL: L08602; AAA16322.1; -.
DR HSSP: P29274; 1MMH.
DR GCRDB: GCR_0525; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PRINTS: PR00242; DOPAMINER.
DR PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE: PS00262; G-PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Lipoprotein; Palmitate.
FT DOMAIN 1 24 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 25 45 1 (POTENTIAL).
FT DOMAIN 46 61 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 62 81 2 (POTENTIAL).
FT DOMAIN 82 98 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 99 120 3 (POTENTIAL).
FT DOMAIN 121 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 164 4 (POTENTIAL).
FT DOMAIN 165 194 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 195 219 5 (POTENTIAL).
FT DOMAIN 220 271 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 272 297 6 (POTENTIAL).
FT DOMAIN 298 310 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 311 330 7 (POTENTIAL).
FT DOMAIN 331 363 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 97 187 BY SIMILARITY.
FT LIPID 345 345 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 363 AA; 40651 MM; 4B47DE240D65DD0 CRC64;

```

```

Query Match 27.4%; Score 493.5; DB 1; Length 363;

```

Best Local Similarity 36.3%; Pred. No. 7.1e-25;
Matches 118; Conservative 59; Mismatches 117; Indels 31; Gaps 8;

```

Oy 21 VRASLYSLAMVLIIFFTLVGNMIVIVYSISHKOLHTP--TNMLHSMVAVDLGLCCAPVPS 79
      || : : ||| ||| ||| : : : : || : : || : : ||| : : ||| : :
Db 23 VRVLGTGFLSVLITSLTGLNLTVCANATKRRLHRSKNTNFVVISLAVSDLLVAVLWPMK 82
Oy 80 MVRSAEHCWYRGEVEYCKIHTSTDMISSASIFHLSTSIDRYAVACDPLRYKAKMNTLVI 139
      || : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : :
Db 83 AVTEVAGFWPFG--AFCCIMAWAFDMCSTASILNLCVYSVDRYAWASSPFYERKMTPRVA 141
Oy 140 CVMFISMSDPAVAFQFMIFELNLFKAEIEYKHYVC-----RGCCSVFSEKISCVLT 193
      ||| : : || : : || : : || : : || : : || : : || : : || : :
Db 142 FVMISGAMTSLVLSF--IPVQLKMHKAQPIGLEVAVASRRDLPTDNCDSLARTVAISS 199
Oy 194 FMSFTYIPGSIIMLCVYRITLIKAEQARLIS--DANOKIQT-----GLEMKNG 239
      || : : ||| : : || : : ||| : : ||| : : ||| : : ||| : :
Db 200 SLISFYIPVALMITYQIYIRIAKOKRIRISALERAESAQIRHDSNGSGSNDLESFK 259
Oy 240 ISQSKERAAVTLGIVNGVFLICMCPCFTIVMDPLHY-----IKPTLNDVLWFGY 293
      || : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : :
Db 260 LSRFRRETKVLKTLTSVINGVFCFCLLPFILLNCWVPFCCKRTISNGLPICSPITPDVFWFGW 319
Oy 294 LNSTFNPAVYAFETPMFRKALKMML 318
      || : : ||| : : ||| : : ||| : : ||| : : ||| : :
Db 320 ANSSLNPITIAFNAD--FRRAFAILL 343

```

| | | |
|---------------|-----------|--------------|
| RESULT | 8 | |
| HH2R_MOUSE | | |
| ID_HH2R_MOUSE | STANDARD; | PRT; 358 AA. |

DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HISTAMINE H2 RECEPTOR (H2R) (GASTRIC RECEPTOR I).
 GN HRH2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/OLA;
 RC MEDLINE=97092891; PubMed=8938453;
 RA Kobayashi T., Inoue I., Jenkins N.A., Gilbert D.J., Copeland N.G.,
 RA Watanabe T.;
 RT "Cloning, RNA expression, and chromosomal location of a mouse
 RT histamine H2 receptor gene";
 RL Genomics 37:390-394(1996).
 CC -1- FUNCTION: THE H2 SUBCLASS OF HISTAMINE RECEPTORS MEDIATES GASTRIC
 CC ACID SECRETION. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G
 CC PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

| | G-protein coupled | receptor: | Transmembrane; Glycoprotein. |
|----|-------------------|-----------|--|
| KM | DOMAIN | 1 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 23 | 1 (POTENTIAL). |
| FT | DOMAIN | 45 | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM | 58 | 2 (POTENTIAL). |
| FT | DOMAIN | 82 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 93 | 3 (POTENTIAL). |
| FT | DOMAIN | 115 | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM | 135 | 4 (POTENTIAL). |
| FT | DOMAIN | 160 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 180 | 5 (POTENTIAL). |
| FT | DOMAIN | 204 | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM | 234 | 6 (POTENTIAL). |
| FT | DOMAIN | 258 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 267 | 7 (POTENTIAL). |
| FT | DOMAIN | 289 | CYTOPLASMIC (POTENTIAL). |
| FT | SITE | 98 | ESSENTIAL FOR HISTAMINE BINDING (BY SIMILARITY). |
| FT | SITE | 185 | ESSENTIAL FOR TIOFIDINE BINDING AND IMPLICATED IN HISTAMINE BINDING (BY SIMILARITY). |
| FT | SITE | 189 | IMPLICATED IN HISTAMINE BINDING (BY SIMILARITY). |
| FT | CARBOHYD | 4 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | DISULFID | 91 | BY SIMILARITY. |
| SO | SEQUENCE | 358 AA; | DDBBAIDTIB6927 CXC64; |

Query Match 27.38; Score 492.5; DB 1; Length 358;

Best Local Similarity 35.0%; Pred. NO. 8.4e-25;
Matches 115; Conservative 54; Mismatches 111; Indels 49; Gaps 7;

```

0Y 30 VLIITITLV-----GNLIVISIHFPOLHPMTMLLHSMATYDFLGLCCIMVYSWRS 84
Db 21 ISVATFLITITVAGNVVYCLAISLNRRLSRLTNCFCIVSLAATFLLLGLIYMFESAIIQL 80
0Y 85 EHCWTFEGVECKHITSDIMLSSASIFHLSITSIDRYXACDPLRYAKKNILIVICMIF 144
Db 81 SFKMFEGGVFNENITSLDMVLCTASILMLNLSIDRCAVATDPLRYIVYLPVRAIASLV 140
0Y 145 ISWSYPAFAECMIFFLELNFAGABEIIYKHWKRCG-----CSYFEKISGVLTFMFTSY 199
Db 141 FLWVISTISLFLSIHLGWSNRNG-----TRGNDITFECKKVQVNEVYGLDGMWTFY 191
0Y 200 IPGSMILCVYRYIYLAAEQARLISDANOKIOLBENKNIGSOSK-----ERKAVKTGI 254
Db 192 LPLMLTCVYTRYIEFLAKREOKAKRI-----NHLSKKAATIEHKRTVYTLAA 237
0Y 255 VMGVFLICMCEFFICTVM-----DPLHYIIPETLNDVLIMESYLNSTFNPNVYAEFP 308
Db 238 VMGAVITVCMPEPFAFVYRGLRGDDP-----VNEVEGIVLMLGYASVSPILLYATLNR 292
0Y 309 WFRKALKMMLEFGKIFOKDSRCKLLELS 337
Db 293 DFRMAVIOOLEPCKLASHNSHKTSLRLNNS 321

```

| | | | |
|----|--|---------|--|
| CC | intermediates, efficiency, deep, fighting, fast, produces, through, control, and | RESULT | 9 |
| CC | Between the Swiss Institute of Bioinformatics and the EMBL Outstation - | H2R_RAT | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | ID | H2R_RAT |
| CC | use by non-profit institutions as long as its content is in no way | AC | P25102; |
| CC | modified and this statement is not removed. Usage by and for commercial | DT | 01-MAY-1992 (Rel. 22, Created) |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | DT | 01-MAY-1992 (Rel. 22, Last sequence update) |
| CC | or send an email to license@isb-sib.ch). | DT | 20-AUG-2001 (Rel. 40, Last annotation update) |
| CC | ----- | DE | HISTAMINE H2 RECEPTOR (H2R) (GASTRIC RECEPTOR I). |
| DR | EMBL; D50096; BAA08792.1; - | GN | HRH2. |
| DR | GCRDb; GCR.1155; - | OS | Rattus norvegicus (Rat). |
| DR | MCD; MGI:108482; Hrh2. | OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| DR | InterPro; IPR000276; GPCR_Rhodopsn. | OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus |
| DR | Pfam; PF00001; 7tm_1; 1 | OX | NCBI_TaxID=10116; |
| DR | PRINTS; PR00237; GPCRHHODPSN. | RN | [1] |
| DR | PRINTS; PR00531; HISTAMINEH2R. | RP | SEQUENCE FROM N. A. |
| DR | PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1. | RX | MEDLINE=92028890; Pubmed=1930186; |
| DR | PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1. | | |

RA Rut M., Traiffort E., Arrang J.-M., Leurs R., Schwartz J.-C.;
 RT "Cloning and tissue expression of a rat histamine H2-receptor gene";
 RL Biochem. Biophys. Res. Commun. 179:1470-1478(1991).
 CC -1- FUNCTION: THE H2 SUBCLASS OF HISTAMINE RECEPTORS MEDIATES GASTRIC
 CC ACID SECRETION. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G
 CC PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: S57565; AAB19935.1; -
 DR PIR: J01278; J01278.
 DR GCRDB: GCR_0308; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PRINTS: PR00531; HISTAMINEH2R.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 22
 FT TRANSMEM 1 22
 FT DOMAIN 23 44
 FT TRANSMEM 45 57
 FT DOMAIN 58 81
 FT TRANSMEM 82 92
 FT DOMAIN 93 114
 FT TRANSMEM 115 134
 FT DOMAIN 135 159
 FT TRANSMEM 160 179
 FT DOMAIN 180 203
 FT TRANSMEM 204 233
 FT DOMAIN 234 257
 FT TRANSMEM 258 266
 FT DOMAIN 267 288
 FT TRANSMEM 289 358
 FT SITE 98
 FT SITE 185
 FT SITE 185
 FT SITE 189
 FT SITE 189
 FT CARBOHYD 4 4
 FT DISULFID 91 173
 FT SEQUENCE 358 AA; 40253 MW; 4886969B7B5DDDC CRC64;
 SQ
 Query Match 27.0%; Score 486; DB 1; Length 358;
 Best Local Similarity 34.0%; Pred. No. 2, 1e-24;
 Matches 117; Conservative 56; Mismatches 123; Indels 48; Gaps 7;

OY 245 -----ERKAVKTLGVNGVFLICWCPFFICTVMDPFLHY-----IIPPLINDVLWFGY 293
 DB 224 ATRREKATVTLVAANVACAFITCFFPTFAV-----YRLRDDDAINEAVEGIVLMTGY 277
 OY 294 LNSTFNDPVYAFEPYPERKALKMMLFGKIPKDKSSRCKFLFELS.337
 DB 278 ANSALNPDIYVAALNRDERTAYQDLFCKFASHSHKTSRLNNS 321
 RESULT 10
 ID DBDR_XENLA STANDARD; PRT; 457 AA.
 AC P42290;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE D(1B) DOPAMINE RECEPTOR (D(5) DOPAMINE RECEPTOR).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95024150; PubMed-7937989;
 RA Sugamori K.S., Demchishyn L.L., Chung M., Niznik H.B.;
 RT "D1A, D1B, and D1C dopamine receptors from Xenopus laevis";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10536-10540(1994).
 CC -1- FUNCTION: THIS IS ONE OF THE FIVE TYPES (D1, D2, D3, D4, D5) OF RECEPTORS
 CC FOR DOPAMINE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G
 CC PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: BRAIN AND KIDNEY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U07864; AAA50829.1; -
 DR HSSP: P29274; 1MMH.
 DR GCRDB: GCR_1321; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PRINTS: PR00242; DOPAMINER.
 DR PRINTS: PR00566; DOPAMINED1B.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein;
 KM Multigene family; Lipoprotein; Palmitate.
 FT DOMAIN 1 41
 FT TRANSMEM 42 67
 FT DOMAIN 68 78
 FT TRANSMEM 79 105
 FT DOMAIN 106 114
 FT TRANSMEM 115 137
 FT DOMAIN 138 156
 FT TRANSMEM 157 181
 FT DOMAIN 182 205
 FT TRANSMEM 206 231
 FT DOMAIN 232 282
 FT TRANSMEM 283 309
 FT DOMAIN 310 326
 FT TRANSMEM 327 351
 FT DOMAIN 352 457
 FT CARBOHYD 24 24
 FT DISULFID 114 199
 FT SEQUENCE 457 AA; 10253 MW; 4886969B7B5DDDC CRC64;
 SQ
 Query Match 27.0%; Score 486; DB 1; Length 358;
 Best Local Similarity 34.0%; Pred. No. 2, 1e-24;
 Matches 117; Conservative 56; Mismatches 123; Indels 48; Gaps 7;


```

RX MEDLINE=96130324; PubMed=8543023;
RA Harwood G.S., Lockyer M., Giles H., Fairweather N.;
RT "Cloning and characterisation of the rabbit 5-HT1D alpha and 5-HT1D
RT beta receptors.";
RN FEBS Lett. 377:73-76(1995).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE;
RA MEDLINE=97032156; PubMed=8878052;
RA Bard J.A., Kucharewicz S.A., Zgombick J.M., Weinschank R.L.,
RA Branchek T.A., Cohen M.L.;
RT "Differences in ligand binding profiles between cloned rabbit and
RT human 5-HT1D alpha and 5-HT1D beta receptors: ketanserin and
RT methiothepin distinguish rabbit 5-HT1D receptor subtypes.";
RL Nauyn Schmiedeberg Arch. Pharmacol. 354:237-244(1995).
CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR 5-
CC HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS
CC AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT INHIBITS ADENYLATE
CC CYCLASE ACTIVITY.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST TO THE OTHER 5HT-1 SUBTYPE RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: 250162; CAA90530.1; -
DR EMBL: 060825; AAB58466.1; -
DR GCRDB: GCR_1293; -
DR GCRDB: GCR_1577; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1.1.
DR PRINTS: PR00237; GPCRHHODOPS.
DR PRINTS: PR00514; 5HT1DRECEPT.
DR PRINTS: PR01101; 5HT1DRECEPT.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1;
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2;
DR KMW G-protein coupled receptor; Transmembrane; Glycoprotein;
DR Multigene family.
FT DOMAIN 1 38 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 39 62 1 (POTENTIAL).
FT DOMAIN 63 75 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 76 98 2 (POTENTIAL).
FT DOMAIN 99 108 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 109 134 3 (POTENTIAL).
FT DOMAIN 135 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 176 4 (POTENTIAL).
FT DOMAIN 177 194 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 195 218 5 (POTENTIAL).
FT DOMAIN 219 302 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 303 326 6 (POTENTIAL).
FT DOMAIN 327 335 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 336 360 7 (POTENTIAL).
FT DOMAIN 361 377 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 111 188 BY SIMILARITY.
SQ SEQUENCE 377 AA; 41500 MW; FC1441678AB82B0A CRC64;

```

```

Query Match 26.4%; Score 475; DB 1; Length 377;
Best Local Similarity 32.0%; Pred. No. 1,le-23;
Matches 112; Conservative 62; Mismatches 120; Indels 56; Gaps 7;

```

```

OY 21 VRASLTLVAVLITLVGNLIVIVSISHFKQLHPTNMLIHSMATVDFLGLGVMPYSW 80

```

```

DB 37 LKSLANVLSITIVAVLSTFVLTTLTRKRLTPANVILGSLATDVLVLPISI 96
OY 81 VRSAEHQWFGVEYFCKIHSTIDIMLSASIFHLSPFISIDRYAVACPLRKARMMILVIC 140
DB 97 AYTHTWNGQVLCIDIWSSDITCCASILHCVIALDRMAITDALEYSKRRTAGHAA 156
OY 141 VMFISVPAVAFPGFIPLFNKAEELYYKHVHCRCGCSFYSIS-GVLTFTMSFY 199
DB 157 AMIAVVAWALSICISIPLEFVR-OAKAHEV-----SDCLVNTQISITVYSTGCAFY 207
OY 200 IPSIMLCVYRYILLAKEO-----ARLSDA----- 226
DB 208 IPSVLIVLXGRYIMARNILNPPSLXGKRTTALHITSASQSLCSLSPSGEGSHS 267
OY 227 -----NOKLQIGLEKNKGISOSKERRAKVTGLGVGLWCPCPFCTVMDPEL 276
DB 268 AGSPLEFNPYRIKLADSVLERKRIISAREKRRKRTTGIIIGAFICGWLPEFVASLVLPIC 327
OY 277 H-YIIPPTLVNVLWGVNLSFTFNPVYAFYFPWRKAKMML-EGKIF 323
DB 328 RDSQWMPGLEDFEFTWGLNLSLNPITYVFNEDFRQAFORVIRPKAR 377
RESULT 13
5HT1D HUMAN STANDARD: PRT: 377 AA.
AC P28221;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 5-HYDROXYTRYPTAMINE 1D RECEPTOR (5-HT-1D) (SEROTONIN RECEPTOR)
DE (5-HT-1D-ALPHA).
OS HTR1D OR HTR1DA.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91342595; PubMed=1652050;
RX Hamblin M.W., Metcalf M.A.;
RT "Primary structure and functional characterization of a human 5-HT1D-
RT type serotonin receptor.";
RL Mol. Pharmacol. 40:143-148(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92228840; PubMed=1565658;
RX Weinschank R.L., Zgombick J.M., Macchi M.J., Branchek T.A.,
RX Hartig P.R.;
RT "Human serotonin 1D receptor is encoded by a subfamily of two
RT distinct genes: 5-HT1D alpha and 5-HT1D beta.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3630-3634(1992).
CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR 5-
CC HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS
CC AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT INHIBITS ADENYLATE
CC CYCLASE ACTIVITY.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST TO THE OTHER 5HT-1 SUBTYPE RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: M89955; AAA35491.1; -
DR EMBL: M81589; AAA60315.1; -

```


QY 133 KNNILVICVNIIFISNVPAPVAFGMIFELNFKGAEIYKHKHCGSGSVFSKTS-CV 191
DB 146 RTAGHAAMIAVAWAISICISIPLEFMR-QATAHEEM-----SDCLVNTSQTSTYT 196
QY 192 LFEWTSFYIPSGIMLCVYRIYLAKEQ-----ARLISDA----- 226
DB 197 YSTGCAFYIPISILLITLYGRIVYAARSRLINPSTLYGKRFTTAQILITGSAGSSLCISLDS 256
QY 227 -----NQ-----KIQIGLEMKNGISQSKERRAVTLGIVMGVFLICMCPFTI 268
DB 257 LHESHTHWGSPLEFQVAKIADSLTERKRISAAREKATPTLGIILCAFIICWLPFV 316
QY 269 CTVMDP-----FLHYIIPPTNDVLIIMFGYLNSTFNPVYAFYFPMFKALMFLGKI 322
DB 317 VSLVLEPICDSCMIH-----PALFDFPTWLGYNLSLINPVIYIVFNEDEFQARQRYVH---- 369
QY 323 FOKDS 327
DB 370 FRKAS 374

RESULT 15
H2R_HUMAN STANDARD: PRT; 359 AA.
AC P25021: Q14464;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1993 (Rel. 22, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HISTAMINE H2 RECEPTOR (H2R) (GASTRIC RECEPTOR 1).
GN HRH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91337087; PubMed=1714721;
RA Gantz I., Munzert G., Tashiro T., Schaeffer M., Wang L.-D.,
RA DelValle J., Yamada T.;
RL "Molecular cloning of the human histamine H2 receptor";
RL Biochem. Biophys. Res. Commun. 178:1386-1392(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=9525318; PubMed=7755641;
RA Nishi T., Koike T., Oka T., Maeda M., Futai M.;
RT "Identification of the promoter region of the human histamine H2-
RT receptor gene.";
RL Biochem. Biophys. Res. Commun. 210:616-623(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=99297624; PubMed=10371214;
RA Murakami H., Sun-Wada G., Matsumoto M., Nishi T., Wada Y., Futai M.;
RT "Human histamine H2 receptor gene: multiple transcription initiation
RT and tissue-specific expressionl.";
RL FEBS Lett. 451:327-331(1999).
RN [4]
RP SEQUENCE OF 4-351 FROM N.A., AND POLYMORPHISM.
RC TISSUE=Brain;
RX MEDLINE=96414531; PubMed=8817552;
RA Orange P.R., Heath P.R., Wright S.R., Pearson R.C.A.;
RT "Allelic variations of the human histamine H2 receptor gene.";
RL NeuroReport 7:1293-1296(1996).
RN [5]
RP REVIEW.
RX MEDLINE=98042107; PubMed=9374694;
RA DelValle J., Gantz I.;
RT "Novel insights into histamine H2 receptor biology.";
RL Am. J. Physiol. 273:G987-G996(1997).
CC -I- FUNCTION: THE H2 SUBCLASS OF HISTAMINE RECEPTORS MEDIATES GASTRIC
CC ACID SECRETION. ALSO APPEARS TO REGULATE GASTROINTESTINAL MOTILITY
CC AND INTESTINAL SECRETION. POSSIBLE ROLE IN REGULATING CELL GROWTH

CC AND DIFFERENTIATION. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY
CC G PROTEIN WHICH ACTIVATE ADENYLYL CYCLASE AND, THROUGH A SEPARATE
CC G PROTEIN-DEPENDENT MECHANISM, THE PHOSPHOLIPID/PROTEIN KINASE
CC (PKC) SIGNALING PATHWAY (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -I- DISEASE: ANTAGONISTS FOR THIS RECEPTOR HAVE PROVEN TO BE EFFECTIVE
CC THERAPY FOR ACID PEPTIC DISORDERS OF THE GASTROINTESTINAL TRACT.
CC CERTAIN ANTAGONISTS ARE USED IN THE TREATMENT OF NEUROPSYCHIATRIC
CC AND NEUROLOGICAL DISEASES SUCH AS SCHIZOPHRENIA, ALZHEIMER'S
CC DISEASE AND PARKINSON'S DISEASE.
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----

DR EMBL: M64799; AAA58647.1; -
DR EMBL: D49783; BAA08618.1; -
DR EMBL: AB023486; BAA84279.1; -
DR EMBL: X98133; CAA66832.1; -
DR PIR: JH0449; JH0449.
DR GCRDB: GCR_0176; -
DR GCRDB: GCR_1836; -
DR GCRDB: GCR_2080; -
DR MIM: 142703; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PRINTS: PR00531; HISTAMINEH2R.
DR PROSITE: PS00237; G-PROTEIN_REC_P1.1; 1.
DR PROSITE: PS50262; G-PROTEIN_REC_P1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Polymorphism.
FT DOMAIN 1 22
FT TRANSMEM 23 44
FT DOMAIN 45 57
FT TRANSMEM 58 81
FT DOMAIN 82 92
FT TRANSMEM 93 114
FT DOMAIN 115 134
FT TRANSMEM 135 159
FT DOMAIN 160 180
FT TRANSMEM 181 204
FT DOMAIN 205 234
FT TRANSMEM 235 258
FT DOMAIN 259 267
FT TRANSMEM 268 289
FT DOMAIN 290 359
FT SITE 98 98
FT SITE 186 186
FT SITE 190 190
FT CARBOHYD 4 4
FT DISULFID 91 174
FT VARIANT 217 217
FT VARIANT 231 231
FT VARIANT 268 268
FT VARIANT 268 268
FT CONFLICT 133 133
FT CONFLICT 175 175
FT CONFLICT 207 207
SQ SEQUENCE 359 AA; 40098 MW; 9835AE2BA60B9B0F CRC64;

Query Match

26.2%; Score 472.5; DB 1; Length 359;

Best Local Similarity 32.68; Pred. No. 1.5e-23;
Matches 106; Conservative 64; Mismatches 124; Indels 31; Gaps 5;

```
OY 18 SNDVRASLYSLMVLILTLVGNLIIVISISHPOLHTPTNMLIHSMTVDLGLGLMP 77
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 14 STACKITTVLAVLILITVAGNVVCLAVGLNRLRLNLTMCFTVSLAITDLGLGLVP 73

OY 78 YSAVSAEHCWYFGEVFCCKIHTSTDIMLSASIFHLSPISIDRYAVCDPLRYKAKNNIL 137
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 FSAIYQLSCKMSFGKVCNITYSIDVMCTASILNLFMISIDRYCAVMDPLRYPLVTPV 133

OY 138 VICVMIFISMSVPAVFAFGMTFLELNFKGAEEIYKHHVCHRGCSVFPSKISGVLTMTS 197
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 134 RVAISLVLIWVISITLSFLSIHLGWNRSRNETS---KGNHTTSKCKVQVNEVYGLVDGLVT 190

OY 198 FYIPGSIMLCVYRRIYLAKEQARLISDANQKLGLEMKNGISQSK-----ERKAVKTL 252
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 191 FYDPLLIMCTIYRIFFKVARDOAKRI-----NHISWKAATIREHKAIVTL 236

OY 253 GIVAGVFLICWCPFFICTVM-----DPELHYIIPPTLNDVLIWFGYLNSTFNPWYAFY 307
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 AAVMGAFILICWFPYETAFAVYGLRGDDAINEV---LEAIYIWLGYANSALNPILYIALN 292

OY 308 PWERKALKMMLFGKIFOKDSSRCKL 332
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 RDFRTGYOQLFCRLANRNSHKTSL 317
```

Search completed: December 11, 2001, 16:10:45
job time: 121 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2001, 16:07:43 ; Search time 12.59 Seconds
(without alignments)
604.139 Million cell updates/sec

Title: US-09-633-145-2

Perfect score: 1801
Sequence: 1 MPFCHNIINISCVKNMNSND.....FGKIFOKDSSRCKLFLFLSS 338

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 813.5 | 45.2 | 343 | 2 | US-08-788-539A-2 |
| 2 | 782 | 43.4 | 338 | 3 | US-09-286-805-4 |
| 3 | 743 | 41.3 | 348 | 3 | US-09-286-805-2 |
| 4 | 677.5 | 37.6 | 337 | 2 | US-08-467-559B-2 |
| 5 | 534 | 29.7 | 388 | 1 | US-08-446-822-8 |
| 6 | 529 | 29.7 | 388 | 5 | PCT-US93-12586-8 |
| 7 | 529 | 29.4 | 387 | 1 | US-07-996-772A-2 |
| 8 | 529 | 29.4 | 387 | 1 | US-08-446-822-2 |
| 9 | 529 | 29.4 | 387 | 5 | PCT-US93-12586-2 |
| 10 | 527 | 29.3 | 406 | 1 | US-07-996-772A-4 |
| 11 | 527 | 29.3 | 406 | 1 | US-08-446-822-4 |
| 12 | 527 | 29.3 | 406 | 5 | PCT-US93-12586-4 |
| 13 | 486 | 27.0 | 358 | 2 | US-08-748-485-6 |
| 14 | 478.5 | 26.6 | 376 | 1 | US-07-817-920-5 |
| 15 | 478.5 | 26.6 | 376 | 1 | US-08-117-006-5 |
| 16 | 478.5 | 26.6 | 376 | 1 | US-08-216-594-5 |
| 17 | 478.5 | 26.6 | 376 | 5 | PCT-US93-00149-5 |
| 18 | 477.5 | 26.5 | 375 | 1 | US-08-370-542-5 |
| 19 | 477.5 | 26.5 | 375 | 1 | US-08-542-358-5 |
| 20 | 477.5 | 26.5 | 375 | 3 | US-09-018-351-5 |
| 21 | 474.5 | 26.3 | 359 | 3 | US-08-875-540-15 |
| 22 | 474 | 26.3 | 377 | 2 | US-08-461-812-2 |
| 23 | 474 | 26.3 | 377 | 2 | US-08-157-185-14 |
| 24 | 474 | 26.3 | 377 | 3 | US-08-281-526B-14 |
| 25 | 472.5 | 26.2 | 359 | 2 | US-08-467-568-13 |
| 26 | 472.5 | 26.2 | 359 | 2 | US-08-748-485-4 |
| 27 | 472.5 | 26.2 | 359 | 2 | US-08-103-170-2 |

| | | | | | | |
|----|-------|------|-----|---|-------------------|-------------------|
| 28 | 472.5 | 26.2 | 359 | 2 | US-09-030-562-13 | Sequence 13, Appl |
| 29 | 461.5 | 25.6 | 348 | 3 | US-08-875-540-13 | Sequence 13, Appl |
| 30 | 459.5 | 25.5 | 359 | 2 | US-08-748-485-5 | Sequence 5, Appl |
| 31 | 458 | 25.4 | 446 | 1 | US-07-626-618A-21 | Sequence 21, Appl |
| 32 | 458 | 25.4 | 446 | 1 | US-08-333-977-21 | Sequence 21, Appl |
| 33 | 456 | 25.3 | 446 | 2 | US-07-969-267B-4 | Sequence 7, Appl |
| 34 | 455 | 25.3 | 483 | 1 | US-08-194-338-7 | Sequence 7, Appl |
| 35 | 454.5 | 25.2 | 477 | 1 | US-07-791-936A-2 | Sequence 2, Appl |
| 36 | 454.5 | 25.2 | 477 | 1 | US-08-383-781B-2 | Sequence 2, Appl |
| 37 | 454.5 | 25.2 | 477 | 2 | US-07-969-267B-2 | Sequence 2, Appl |
| 38 | 452 | 25.1 | 417 | 5 | PCT-US91-00909-2 | Sequence 2, Appl |
| 39 | 452 | 25.1 | 418 | 1 | US-08-194-338-9 | Sequence 9, Appl |
| 40 | 452 | 25.1 | 418 | 3 | US-08-817-869-12 | Sequence 12, Appl |
| 41 | 450.5 | 25.0 | 359 | 1 | US-07-996-772A-10 | Sequence 10, Appl |
| 42 | 450.5 | 25.0 | 359 | 2 | US-08-748-485-3 | Sequence 3, Appl |
| 43 | 450 | 25.0 | 475 | 1 | US-07-686-591-4 | Sequence 4, Appl |
| 44 | 450 | 25.0 | 475 | 1 | US-07-970-715-4 | Sequence 4, Appl |
| 45 | 449.5 | 25.0 | 446 | 1 | US-07-626-618A-22 | Sequence 22, Appl |

ALIGNMENTS

RESULT 1
US-08-788-539A-2
: Sequence 2, Application US/08788539A
: Patent No. 5871967
: GENERAL INFORMATION:
: APPLICANT: SmithKline Beecham Corporation
: TITLE OF INVENTION: CLONING OF A NOVEL G-PROTEIN
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY: USA
: ZIP: 19406
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/788,539A
: FILING DATE: 24-JAN-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Han, William T
: REGISTRATION NUMBER: 34,344
: REFERENCE/DOCKET NUMBER: ATG50047
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-5219
: TELEFAX: 610-270-4026
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 343 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-788-539A-2

Query Match 45.2%; Score 813.5; DB 2: CompuGen 343;
Best Local Similarity 46.0%; Pred. No. 1.9e-63;
Matches 159; Conservative 57; Mismatches 99; Gaps 5;


```

? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: STEFFE, ERIC K
? REGISTRATION NUMBER: 36,688
? REFERENCE/DOCKET NUMBER: 1488, 08400000
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 371-2600
? TELEFAX: (202) 371-2540
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 337 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? OS-08-467-559B-2

```

| | | | | |
|-----------------------|-----------------|-------------------|-----------|------------|
| Query Match | 37.6% | Score 677.5 | DB 2 | Length 337 |
| Best Local Similarity | 39.0% | Pred. No. 1.2e-51 | | |
| Matches 130 | Conservative 67 | Mismatches 125 | Indels 11 | Gaps 4 |

```

QY 3 ECHHINIISCKNNKMSNDVRASISLWVLIILTTNVLNIVYSISHPKOLHPPTNLIH 62
Dh 16 FCYO-VNGSCRTHTGILQIVLYITFCAGMLIIVLGNVFAFVAFKALHPPTNLLL 74
QY 63 SMATVDFLLGLVMPYSWVRBAEHCWFEFGEFCKIHNSTDIMLSASIFHLSFDIDRY 122
Dh 75 SLADMFELGLVLPSTIRKSVESCEPFGBDLCHMYLDTLFCUTSIFHLCFSDIRHC 134
QY 123 AVCDDPLRYKAMNIIIVICMIFISWSPVAFVAFQCMIFLELNFKAEBEYVKNVHRCGS 182
Dh 135 AICPDLTPYSKFETVRVALRYTLAAGVPAATSTLELYTDVETRISQ-WLEEMPCVSCQ 193
QY 183 VEFKISGVLFEMHSFYIPGSIIMCVYYRYIILNKBDQARLISDANKLOIGLEBKNIISQ 242
Dh 194 LLLKFMWMLNF-PLFVFPCLIMISLKYKIRFVATROAQOITTLISKSL-----AGAA 244
QY 243 SKERRAVKTLGIVNGVPLICSCPFICITVMDPRFLHYIIPRLYNDVLLMFGLYNTNNPV 302
Dh 245 KHERKAATLGIVGYILCLWLPRTIDMDSLHFTIRPLVDFIPLMPAVFNACAPII 304
QY 303 YAFEPYPRKALMMLEGGIKIQOKSSCKCLEFLE 335
Dh 305 YVESIQWERRKALKLTLISOKVESPQTRVDYIOE 337

```

RESULT 5
 US-08-446-822-8
 Sequence 8, Application US/08446822
 Patent No. 5766879
 GENERAL INFORMATION:
 APPLICANT: SYNATIC PHARMACEUTICAL CORPORATION
 TITLE OF INVENTION: DNA ENCODING 5-HT4 SEROTONIN RECEPTORS
 TITLE OF INVENTION: AND USES THEREOF
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: COOPER & DUNHAM
 STREET: 30 ROCKEFELLER PLAZA
 CITY: NEW YORK
 STATE: NEW YORK
 ZIP: 10112
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,822
 FILING DATE: June 1, 1995
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: White, P., John
 REGISTRATION NUMBER: 28,678

```

? REFERENCE/DOCKET NUMBER: 42667-A-PCT-US/JPW/MA
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 278-0400
? TELEFAX: (212) 391-0525
? INFORMATION FOR SEO ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 388 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? OS-08-446-822-8

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 29.7%; | Score 534; | DB 1; | Length 388; |
| Best Local Similarity | 37.4%; | Pred. No. 4.4e-39; | | |
| Matches 116; | Conservative 53; | Mismatches 121; | Indels 20; | Gaps 5; |

| | | | |
|----|-----|--|-----|
| Qy | 25 | LXSLMVLITLTLGNLIVIVISHSFKOLH-TTPNNMLHSMAVDFLLCGLWPEYSVR | 83 |
| Db | 22 | LLTFLSVYILMALLGNLLVMVAWCMQOLKRIKTNFYISLAFNDDLVSVLWPEGAIEL | 81 |
| Qy | 84 | AEHCWYGEVCEVCKLHTSTIDMLSSASIFHLSEFISIDRYAV-CDPLKAKAKMLVLCVM | 142 |
| Db | 82 | VQDITWYGEVCEVCLRTSIDVLLTTLTASIFHLCSILDRYALCCDPLVRYNNKMPRLALM | 141 |
| Qy | 143 | IFISWYVAEVAFGMIFELNFKGAEEIYUKHVHCRGG---CSVEFSKISGVLTMTSF | 198 |
| Db | 142 | LGGCMVIFPTFISFLPMQGMNITIDILIEKRKNQMSNSTYCVFMYNKKPYALCSVAF | 201 |
| Qy | 199 | YIPGSIIMLCVYRYIYLLAKEQARLISDANKLQIGLEMKNGISG-----KPKKA | 248 |
| Db | 202 | YIPLFLMWLAYRYIYVAKAEHAHOI---QKLOKAGASSSESRPASDOHSTHRMRTETKA | 257 |
| Qy | 249 | VKTIGIIVAGVLIQWCEFFLWADPFLHYLIPTLNDVLIWEGYSLNSTPNVYAEFP | 308 |
| Db | 258 | AKTICIIMGCGCLCMAPFEVYINVDPIFYDVPGQVMTAFMLGIVIRSGINPFLYAFLNK | 314 |
| Qy | 309 | WFKRAKAKML 318 | |
| Db | 318 | SFRRAFLITL 327 | |

RESULT 6
PCT-US93-12586-8
Sequence 8, Application PC/TUS9312586
GENERAL INFORMATION:
APPLICANT: SYNAPTEC PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: DNA ENCODING 5-HT4 SEROTONIN RECEPTORS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM
STREET: 30 ROCKEFELLER PLAZA
CITY: NEW YORK
STATE: NEW YORK
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12586
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White, P., John
REGISTRATION NUMBER: 28, 678
REFERENCE/DOCKET NUMBER: 42667-A-PCT/JPW/TEP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-12586-8

Query Match 29.7%; Score 534; DB 5; Length 388;
Best Local Similarity 37.4%; Pred. No. 4,4e-39;
Matches 116; Conservative 53; Mismatches 121; Indels 20; Gaps 5;

QY 25 LYSIMALLITLVGNLYIVISISHKOLH-TPTNMLHSMATVDLLGCVMPYMSVRS 83
DB 22 LTFEFTVLTMLLGNLVAVACMDROLKIKTNFVLSLAADLLSVLVNAFGAIEL 81
QY 84 AEHCWYGEVEFCIKHTSTDIMLSSASIFHLSTIDRYAV-CDPLRYAKNMILVYCV 142
DB 82 VQDIWYGEVFLVRLSDVLTFTASIFHLCCISLDRYAICCPLYRRKMTPLAIALM 141
QY 143 IFISWVPAVFAFGMTLELNFKAEEIYKHHVCRGG---CSVEFSKISGVLTMTSE 198
DB 142 LGGCWVTFPISEFLPIMGWNNIGIIDLIEKRKNQNSSTYCVFWNKPAYATCSVAF 201
QY 199 YIPGSIIMLCVYRYIYLAIEQARLISDANKQIDGLEMKNGISQ-----KERK 248
DB 202 YIPFLMLAVLYRYIYTAEEHAQI---QMLQACASSESPQSDOHSHTRMRTETRA 257
QY 249 VTLGIVMGVFLICMCPFTCTVMDPELHYIIPTLNDVLMFGYLNSTFNPVAFYFP 308
DB 258 AATLCITMGCFCLCMAPFEVTNIVDPFIDYTVPGQWTAFLMIGYINSGLNPLVAF LK 317
QY 309 WPKKALKMML 318
DB 318 SERRAFLIIL 327

RESULT 7

US-07-996-772A-2
Sequence 2, Application US/07996772A
Patent No. 5472866
GENERAL INFORMATION:
APPLICANT: Gerald, Christophe
APPLICANT: Hartig, Paul R.
APPLICANT: Branchek, Theresa A.
APPLICANT: Weinschenk, Richard L.
TITLE OF INVENTION: DNA ENCODING 5-HT4A SEROTONIN
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM
STREET: 30 ROCKEFELLER PLAZA
CITY: NEW YORK
STATE: NEW YORK
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/996,772A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, P., John
REGISTRATION NUMBER: 28, 678
REFERENCE/DOCKET NUMBER: 42667/JPM/TEP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-996-772A-2

Query Match 29.4%; Score 529; DB 1; Length 387;
Best Local Similarity 35.4%; Pred. No. 1.2e-38;
Matches 111; Conservative 61; Mismatches 114; Indels 28; Gaps 6;

QY 25 LYSIMALLITLVGNLYIVISISHKOLH-TPTNMLHSMATVDLLGCVMPYMSVRS 83
DB 22 LTFEFTVLTMLLGNLVAVACMDROLKIKTNFVLSLAADLLSVLVNAFGAIEL 81
QY 84 AEHCWYGEVEFCIKHTSTDIMLSSASIFHLSTIDRYAV-CDPLRYAKNMILVYCV 142
DB 82 VQDIWYGEVFLVRLSDVLTFTASIFHLCCISLDRYAICCPLYRRKMTPLAIALM 141
QY 143 IFISWVPAVFAFGMTLELNFKAEEIY---YKHVCRGGCSVEFSKISGVLTMTSE 198
DB 142 LGGCWVTFPISEFLPIMGWNNIGIYDIEKRKNHNSSTFCVFWNKPAYATCSVAF 201
QY 199 YIPGSIIMLCVYRYIYLAIEQARLISDANKQIDGLEMKNGISQK----- 244
DB 202 YIPFLMLAVLYRYIYTAEEHA-----QQLQACASSESPQSDOHSHTRMRT 253
QY 245 ERKAVTGLIVMGVFLICMCPFTCTVMDPELHYIIPTLNDVLMFGYLNSTFNPVAF 304
DB 254 ETKAATLIVCMGCFCLCMAPFEVTNIVDPFIDYTVPEKWTAFMLGYNGLNPLVA 313
QY 305 FEYMPERKALKMML 318
DB 314 FLNKSERRAFLIIL 327

RESULT 8

US-08-446-822-2
Sequence 2, Application US/08446822
Patent No. 5766879
GENERAL INFORMATION:
APPLICANT: SYNAPTIC PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: DNA ENCODING 5-HT4 SEROTONIN RECEPTORS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM
STREET: 30 ROCKEFELLER PLAZA
CITY: NEW YORK
STATE: NEW YORK
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,822
FILING DATE: June 1, 1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: White, P., John
REGISTRATION NUMBER: 28, 678
REFERENCE/DOCKET NUMBER: 42667-A-PCT-US/JPM/MAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid

| | | | | |
|-----------------------|-----------------|--|-----------|------------|
| Query Match | 29.3% | Score 527 | DB 1 | Length 406 |
| Best Local Similarity | 35.0% | Pred. No. 1.9e-38 | | |
| Matches 110 | Conservative 62 | Mismatches 114 | Indels 28 | Gaps |
| QY | 25 | LYSLMVLITLTTLVGNLIVIVSISHFKQLH-TTPNNMLIHSNATVDFLGLGLVMPYSWRS | 83 | |
| Db | 22 | LTLEFAYVILMAILGNLLVMVAVCGRDQRLKRIKTNFYISLAFADLLSVLYVNAFGAIEL | 81 | |
| QY | 84 | AERHWYGEVFCFKHTHTSDIMLSASIFPHLSFISIDRYAV-CDPLRYKAKMNLIVICM | 142 | |
| Db | 82 | VQDIMPFGEMECIATRSIDVLTLTASIFHCLCSLDRYVAICCPLYRNKMFPLRIALM | 144 | |
| QY | 143 | IFISMSYPAVAFQMIFPLELNFKAEEIY---KKHHCHRGCGCVFFSKTISGVLTMTSF | 198 | |
| Db | 142 | LGGCWVITPMFISPLIPMOGMNNGISIVYTEKRKNHNSNSTFCVEMNKKPYALTICSVAF | 201 | |
| QY | 199 | YIPGSIIMLCVYRYRIYLAKEQARLISDANOKLQJGLEMKNGNISOSK----- | 244 | |
| Db | 202 | YIPLFLVAVLARIYRYVAKHEA-----QOIQM-LDRAGCTSSRSKROTADQHSHTHMR | 25 | |
| QY | 245 | ERRAKVTLGLVMGVELICWCPFCITCYVMDPLFLAYIIPPTLNDVLIIMEGYLNSFTFNMYA | 304 | |
| Db | 254 | ETKAKTLCVIMGCEFCQMAPEFVNTIVDPIDYTBKEKVTATLMLGYINSGLNPFLYA | 313 | |
| QY | 305 | FIFYPMERKALKMML | 318 | |
| Db | 314 | FLNKSFRRAFLIIL | 327 | |

RESULT 11
US-08-446-822-4
Sequence 4, Application US/08446822

GENERAL INFORMATION:
APPLICANT: SYNTACTIC PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: DNA ENCODING 5-HT7A SEROTONIN RECEPTORS
TITLE OF INVENTION: DNA ENCODING 5-HT7A SEROTONIN RECEPTORS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM
STREET: 30 ROCKEFELLER PLAZA
CITY: NEW YORK
STATE: NEW YORK
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,822
FILING DATE: June 1, 1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: White, P., John
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 42667-A-PCT-US/JPM/MAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 406 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-822-4

| | | | | |
|---------------------------|--------|--------------------|------------|-------------|
| Query Match | 29.3%; | Score 527; | DB 1; | Length 406; |
| Best Local Similarity | 35.0%; | Pred. No. 1.9e-38; | | |
| Matches 110; Conservative | 62; | Mismatches 114; | Indels 28; | Gaps 6 |

[illegible]

RESULT 12
PCT-US93-12586-4
; Sequence 4, Application PC/TUS9312586

```

APPLICANT: SYNAPTIC PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: DNA ENCODING 5-HT4 SEROTONIN RECEPTORS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESS: COOPER & DUNHAM
STREET: 30 ROCKEFELLER PLAZA
CITY: NEW YORK
STATE: NEW YORK
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12586
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White, P., John
REGISTRATION NUMBER: 28, 678
REFERENCE/DOCKET NUMBER: 42667-A-PCT/JPW/TEP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 406 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-12586-4

```

[illegible]

[illegible]

```

RESULT 13
US-08-748-485-6
: Sequence 6, Application US/08748485
: Patent No. 5817480
: GENERAL INFORMATION:
: APPLICANT: Au-Young, Janice
: APPLICANT: Guegler, Karl J.
: APPLICANT: Goli, Surya K.
: APPLICANT: Moriy, Lynn E.
: TITLE OF INVENTION: NOVEL HISTAMINE H2 RECEPTOR
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: US
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/748,485
: FILING DATE: Herewith
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0159 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 358 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 236184
: US-08-748-485-6

```

Best Local Similarity 34.0%; Pred. No. 6e-35;
Matches 117; Conservative 56; Mismatches 123; Indels 48; Gaps 7.

RESULT 14
 US-07-817-920-5
 Sequence 5, Application US/07817920
 Patent No. 5360735
 GENERAL INFORMATION:
 APPLICANT: Weinschank, Richard L
 APPLICANT: Weinschank, Theresa
 APPLICANT: Hartly, Paul R
 TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR AND
 TITLE OF INVENTION: USES THEREOF
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Cooper & Dunham
 STREET: 30 Rockefeller Plaza
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10112
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/817, 920
 FILING DATE: 19920108
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 1795/39318
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-977-9550
 TELEFAX: 212-664-0525
 TELEEX: 422523 COOP UI
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 376 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal

IMMEDIATE SOURCE:
CLONE: 5-HT1DA
US-07-817-920-5

Query Match 26.6%; Score 478.5; DB 1; Length 376;
Best Local Similarity 34.0%; Pred. No. 2.9e-34;
Matches 122; Conservative 55; Mismatches 111; Indels 71; Gaps 11;

QY 21 VRASLSLNAVLIITLVGNLIYVSIHFKQLHTPTNMLHSMATVDELGLVMPYSN 80
DB 37 LKISLPEVLISVITLATVLSNAFVLTITLLTRKLTNPANLYLSLATDVLVSILVMPISM 96
QY 81 VNSAEHCWYEGVEFKIHTSTIDIMLSASIFHLFSIDRYAVCPRLRYKAKMNIIVIC 140
DB 97 ATTITHTWNEGQILCDIMLSSDITCTASTILHCVIALDRYMAITALEYSKRRTAGHAA 156
QY 141 VMIFISWVPAVAFGMIFLEINFKGADEIYKHHVCRGCSVFESKIS-GVLTFTMTSEY 199
DB 157 TMAIYMAISICISIPPLFWROE-KAQDEM-----SDCLVNTSQISYTIYSTCGARY 207
QY 200 IFSIMLCVYRYLYLAKQO-----ARLIDA-----NOKLQIG----- 233
DB 208 IPSVLLIILYGRILYRAARRILNPPSLSGKRTTAHLITGSAGSVCSLSLSLHEGSHSA 267
QY 234 -----LEMKNGISQSKERKAVKTGIWGVFLICWCPFFICTVMDP-- 274
DB 268 GSPLEFNHVKIKLADSALEKRR-ISAAREKAKIKIIGIITGAFITCMLPEFVVSILVLPIC 326
QY 275 ---FLHYIIPPTLNDVLIWFGYLNSTFNPVYAFEPYPMFRKALKMMEFGKI--FOKDS 327
DB 327 RDSWTH----PGLFDFFTWLGYNLSLNPDIITYVNEERQA-----FOKIYVPRKAS 376

RESULT 15
US-08-117-006-5
Sequence 5, Application US/08117006
Patent No. 5639652

GENERAL INFORMATION:

APPLICANT: Weinshank, Richard L.
APPLICANT: Branchek, Theresa
TITLE OF INVENTION: DATA ENCODING A HUMAN 5-HT1F RECEPTOR AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect, Version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/117,006

FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 1795/39318
TELEPHONE: 212-977-9550
TELEFAX: 212-664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid

STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
IMMEDIATE SOURCE:
CLONE: 5-HT1DA
US-08-117-006-5

Query Match 26.6%; Score 478.5; DB 1; Length 376;
Best Local Similarity 34.0%; Pred. No. 2.9e-34;
Matches 122; Conservative 55; Mismatches 111; Indels 71; Gaps 11;

QY 21 VRASLSLNAVLIITLVGNLIYVSIHFKQLHTPTNMLHSMATVDELGLVMPYSN 80
DB 37 LKISLPEVLISVITLATVLSNAFVLTITLLTRKLTNPANLYLSLATDVLVSILVMPISM 96
QY 81 VNSAEHCWYEGVEFKIHTSTIDIMLSASIFHLFSIDRYAVCPRLRYKAKMNIIVIC 140
DB 97 ATTITHTWNEGQILCDIMLSSDITCTASTILHCVIALDRYMAITALEYSKRRTAGHAA 156
QY 141 VMIFISWVPAVAFGMIFLEINFKGADEIYKHHVCRGCSVFESKIS-GVLTFTMTSEY 199
DB 157 TMAIYMAISICISIPPLFWROE-KAQDEM-----SDCLVNTSQISYTIYSTCGARY 207
QY 200 IFSIMLCVYRYLYLAKQO-----ARLIDA-----NOKLQIG----- 233
DB 208 IPSVLLIILYGRILYRAARRILNPPSLSGKRTTAHLITGSAGSVCSLSLSLHEGSHSA 267
QY 234 -----LEMKNGISQSKERKAVKTGIWGVFLICWCPFFICTVMDP-- 274
DB 268 GSPLEFNHVKIKLADSALEKRR-ISAAREKAKIKIIGIITGAFITCMLPEFVVSILVLPIC 326
QY 275 ---FLHYIIPPTLNDVLIWFGYLNSTFNPVYAFEPYPMFRKALKMMEFGKI--FOKDS 327
DB 327 RDSWTH----PGLFDFFTWLGYNLSLNPDIITYVNEERQA-----FOKIYVPRKAS 376

Search completed: December 11, 2001, 16:08:59
Job time: 76 sec

2001-12-11 16:12:41

2001-12-11 16:12:41

THIS PAGE BLANK (USPTO)